

STIC-Biotech/ChemLib

158263

me

From: Slobodyansky, Elizabeth
Sent: Tuesday, July 05, 2005 7:58 PM
To: STIC-Biotech/ChemLib
Subject: 10/758,672 SEQ

Please search for case 10/758,672:

88

SEQ ID NOs: 1 and 2 against commercial ~~and interference~~ databases.

Thank you.

Elizabeth Slobodyansky, PhD

Primary Examiner
Art Unit 1652
REM 2D83
571-272-0941
MAILBOX 2C70

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2-_____
Date Searcher Picked up: _____
Date Completed: _____
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA#: _____ AA#: _____
Interference: _____ SPDI: _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure#: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other(Specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 8, 2005, 18:41:46 ; Search time 25839 Seconds
(without alignments)
11829.231 Million cell up

Title: US-10-758-672A-1
 Perfect score: 6308
 Sequence: 1 gccaaagattcggcacgagg.....atttttgatttttggtgtttt 6308

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:*

1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_v:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query #			Length	DB	ID	Description
	Score	Match					
1	6308	100.0	6308	6	AR487554	AR487554	Sequence
2	5250	83.2	5250	6	AY061886	AY061886	Homo sapi
3	5136	81.4	5205	6	AK0487565	AK0487565	Sequence
4	5085.6	80.6	5128	9	AF525401	AF525401	Homo sapi
5	4435.4	70.3	6395	6	AR030784	AR030784	Sequence
6	4435.4	70.3	6395	6	AR121463	AR121463	Sequence
7	4435.4	70.3	6395	6	BD245011	BD245011	Nucleic a
8	4435.4	70.3	6395	10	AF061555	AF061555	Mus muscu
9	3473.6	55.1	4529	6	CQ728565	CQ728565	Sequence
10	3055.8	48.4	3059	6	AK714232	AK714232	Sequence
11	3055.8	48.4	3059	6	AK056441	AK056441	Homo sapi
12	2546.8	40.4	2550	6	BD156870	BD156870	Primer fo
13	2546.8	40.4	2550	6	AK877826	AK877826	Sequence
14	2546.8	40.4	2550	9	AK027803	AK027803	Homo sapi
15	2487.8	39.4	4576	9	HS0808548	EX648400	Homo sapi
16	1529.2	24.2	3598	6	C0850354	C0850354	Sequence
17	1529.2	24.2	3598	9	AK127478	AK127478	Homo sapi
18	1381.8	21.9	6300	6	AR487555	AR487555	Sequence
19	1378.6	21.9	5268	9	AY061884	AY061884	Homo sapi

20	1233.6	19.6	6089	6	AR487556	AR487556	Sequence
21	1233.6	19.6	7511	10	BC073642	BC073642	Mus muscu
22	1233.4	19.6	5265	10	AY061885	AY061885	Mus muscu
23	1217.4	19.3	5438	10	AY280958	AY280958	Mus muscu
24	996.2	15.8	1001	6	AR030785	AR030785	Sequence
25	996.2	15.8	1001	6	AR121464	AR121464	Sequence
26	996.2	15.8	1001	6	BD245012	BD245012	Nucleic a
27	994.2	15.8	999	9	AF061556	AF061556	Homo sapi
28	981	15.6	6158	9	AB002347	AB002347	Homo sapi
29	870.2	13.8	6521	10	AK122254	AK122254	Human mRN
30	858	13.6	1330	5	CR387019	CR387019	Mus muscu
31	813.6	12.9	3327	5	CQ755169	CQ755169	Gallus ga
32	813.6	12.9	3502	6	AX965629	AX965629	Sequence
33	797.2	12.6	4208	6	CQ842877	CQ842877	Sequence
34	797.2	12.6	4208	9	AK125795	AK125795	Homo sapi
35	772.8	12.3	3502	10	BC031403	BC031403	Mus muscu
36	764.2	12.1	818	9	HSX521	HSX521	Human parti
37	656.8	10.4	733	6	BD221289	BD221289	Human gen
38	651.6	10.3	2958	9	AK026998	AK026998	Homo sapi
39	647	10.3	712	6	BD149613	BD149613	Primer fo
40	647	10.3	712	6	AX869551	AX869551	Sequence
41	638.5	10.1	756	6	BD221205	BD221205	Human gen
42	515.2	8.2	2475	10	BC026391	BC026391	Mus muscu
43	503.8	8.0	166518	9	AC090514	AC090514	Homo sapi
44	487.8	7.7	2512	9	AK026948	AK026948	Homo sapi
45	392.4	6.2	2518	9	BC064512	BC064512	Homo sapi

ALIGNMENTS

RESULT 1
 AR487554 LOCUS
 DEFINITION
 AR487554 ACCESSION
 AR487554.1 GI:47252804
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Unknown.
 Unclassified.
 1 (bases 1 to 6308)
 Han.H.-O. and Kwak,K.
 Human E3.alpha. ubiquitin ligase family
 Patent: US 6706505-A 1 16-MAR-2004;
 Location/Qualifiers
 1..6308
 /organism="unknown"
 /mol_ttype="genomic DNA"
 source

/mol_type= genomic DNA									
ORIGIN		Query Match	100.0%;	Score 6308;	DB 6;	Length 6308;			
		Best Local Similarity	100.0%;	Pred. No. 0;					
		Matches 6308;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps
QY	1	GCCAAGAATTCCGCACGAGGGGAAAAGCTGAGCCACAGGAACCCAAATTACTTGGCTTTACCT	60						
Db	1	GCCAAGAATTCCGCACGAGGGGAAAAGCTGAGCCACAGGAACCCAAATTACTTGGCTTTACCT	60						
QY	61	CATTGTGTAAGACAAGCGTCAAAAACAGCTTCAACCTATCTTGGAACAGAGAACTTTACCT	120						
Db	61	CATTGTGTAAGACAAGCGTCAAAAACAGCTTCAACCTATCTTGGAACAGAGAACTTTACCT	120						
QY	121	CCAAAGGCTTATCATCTGCTTCCACTTATCCAAACAGCTGCTATGGCCACCTGCCTGTGC	180						
Db	121	CCAAAGGCTTATCATCTGCTTCCACTTATCCAAACAGCTGCTATGGCCACCTGCCTGTGC	180						
QY	181	CGCACTCGAAACACCGCCAGCCGCCACTACTGCTCCACTACACCTGGTGTCTCCCAACCT	240						
Db	181	CGCACTCGAAACACCGCCAGCCGCCACTACTGCTCCACTACACCTGGTGTCTCCCAACCT	240						
QY	241	GATCAGCTGCTTCTGCTGCCATCTTATCCGCTTCTGCCTTCTGCTGAGTAAGTATACA	300						

Db 241 GATCAGCTGCTTGCTGCCATCTTATCCGCTTCTGCTGCTTCTGAGTAAATGTATACA 300
Qy
Db 301 CACCTTGAACACCAATCTACTTTCTGTGCTCTATGAATTTGACTACTCTAGCTGGATC 360
Db 301 CACCTTGAACACCAATCTACTTTCTGTGCTCTATGAATTTGACTACTCTAGCTGGATC 360
Qy 361 CCGAGCTTTTGTGTACACATGTGCAAGTGGCCACGGGGTGAATCCTTAAATAATAGAGAT 420
Db 361 CCGAGCTTTTGTGTACACATGTGCAAGTGGCCACGGGGTGAATCCTTAAATAATAGAGAT 420
Qy 421 GTATGCAACAGTTCCAGCACCAAAACCCAGATATACAAACCAATTCAGCTTACCAAGAGTAC 480
Db 421 GTATGCAACAGTTCCAGCACCAAAACCCAGATATACAAACCAATTCAGCTTACCAAGAGTAC 480
Qy 481 GCCTGATAAATTTAGAGGGGAAAAAATCTCCAGTCCCTTCAGCTCGTGAACGCTTGC 540
Db 481 GCCTGATAAATTTAGAGGGGAAAAAATCTCCAGTCCCTTCAGCTCGTGAACGCTTGC 540
Qy 541 TCCGGGAAGCGGCGGAAGCCACTCTCGAGTCTCGCTCAAAACCCGACTTTCAGGGGCG 600
Db 541 TCCGGGAAGCGGCGGAAGCCACTCTCGAGTCTCGCTCAAAACCCGACTTTCAGGGGCG 600
Qy 601 TCGTAAAAAGTGTGCTCCCTGCTCTCCGACCGGCCACAGGTTCCGCTTGCCTCTGGCCG 660
Db 601 TCGTAAAAAGTGTGCTCCCTGCTCTCCGACCGGCCACAGGTTCCGCTTGCCTCTGGCCG 660
Qy 661 GGGGTCCGGAACCTGACGGGCTCAGTTTCCCTCAAGATGGCGGACGAGGAGCTGGAGTA 720
Db 661 GGGGTCCGGAACCTGACGGGCTCAGTTTCCCTCAAGATGGCGGACGAGGAGCTGGAGTA 720
Qy 721 CTGAGAGATGGAATCAGCGCGAGTTACCCAGACCCCTCAGCGTCTGGCATCTTGGT 780
Db 721 CTGAGAGATGGAATCAGCGCGAGTTACCCAGACCCCTCAGCGTCTGGCATCTTGGT 780
Qy 781 GGGATCAGCAAGTTGAATTTTATCTGCTTTCTGTCATCATTTGGSCAAATTTGGTCCAG 840
Db 781 GGGATCAGCAAGTTGAATTTTATCTGCTTTCTGTCATCATTTGGSCAAATTTGGTCCAG 840
Qy 841 AAATTTACTTTGCTGAAATGGAACCCAGACTTGGAAAAAGCAGGAGAAAGTGTAACAATGT 900
Db 841 AAATTTACTTTGCTGAAATGGAACCCAGACTTGGAAAAAGCAGGAGAAAGTGTAACAATGT 900
Qy 901 CAATATTCACCTCCATGGTACTTATTTGGAGAGATCCAGATATTTGCTTTAGAGA 960
Db 901 CAATATTCACCTCCATGGTACTTATTTGGAGAGATCCAGATATTTGCTTTAGAGA 960
Qy 961 AATTGAAGCACAGTGGAGCATTTACGCTTTTGGGAGGGTTTCAAAAGTGGAGACAA 1020
Db 961 AATTGAAGCACAGTGGAGCATTTACGCTTTTGGGAGGGTTTCAAAAGTGGAGACAA 1020
Qy 1021 CCTATTCTTGGAGGATTTGCAATTGATCCAAATGTGATCTGTATGGACTGCTTCC 1080
Db 1021 CCTATTCTTGGAGGATTTGCAATTGATCCAAATGTGATCTGTATGGACTGCTTCC 1080
Qy 1081 AGGACAGTGTTCATAAAATCATCGTTACAAGATGCATCTTCTACTGGAGGGTCT 1140
Db 1081 AGGACAGTGTTCATAAAATCATCGTTACAAGATGCATCTTCTACTGGAGGGTCT 1140
Qy 1141 GTGACTGTGGAGACACAGAGGCATGGAATACTGGCCCTTTTGTGTAAATCATGAACCTG 1200
Db 1141 GTGACTGTGGAGACACAGAGGCATGGAATACTGGCCCTTTTGTGTAAATCATGAACCTG 1200
Qy 1201 GAAGAGCGGTACTATAAAGAGAAATTCACGCTGCTCGTTGAATGAAGAGTAAATGTGCC 1260
Db 1201 GAAGAGCGGTACTATAAAGAGAAATTCACGCTGCTCGTTGAATGAAGAGTAAATGTGCC 1260
Qy 1261 AAGCAGGAAATATTTCTTCCAGTGAATAAATATGTCTGAGAAATGACTATATGGGAG 1320
Db 1261 AAGCAGGAAATATTTCTTCCAGTGAATAAATATGTCTGAGAAATGACTATATGGGAG 1320
Qy 1321 AGGAAAAAGAACTGCTCTCGAATCCAGATGAAGGAGAAAAATGAAAGATACTATTGTG 1380
Db 1321 AGGAAAAAGAACTGCTCTCGAATCCAGATGAAGGAGAAAAATGAAAGATACTATTGTG 1380

Qy 1381 TCCCTTTCAATGATGAACACCAATTCATATGACACGCTCATATACAGCTTACAAAGAGCTC 1440
Db 1381 TCCCTTTCAATGATGAACACCAATTCATATGACACGCTCATATACAGCTTACAAAGAGCTC 1440
Qy 1441 TTGACTGTGAGCTCGCAGAGGCCAGTTGTCATACCACTGCCATTGACAAAGAGGGTGGTC 1500
Db 1441 TTGACTGTGAGCTCGCAGAGGCCAGTTGTCATACCACTGCCATTGACAAAGAGGGTGGTC 1500
Qy 1501 GGGCTGTAAAGCGGGAGCTTATGCTGCTTCCAGGAGCAAGGAAGATATAAAGAGTC 1560
Db 1501 GGGCTGTAAAGCGGGAGCTTATGCTGCTTCCAGGAGCAAGGAAGATATAAAGAGTC 1560
Qy 1561 ATTACAGAAAATGTCTCTCAACATCCACTCATGTAGAAATATTAACCTCAGAGATTTAGG 1620
Db 1561 ATTACAGAAAATGTCTCTCAACATCCACTCATGTAGAAATATTAACCTCAGAGATTTAGG 1620
Qy 1621 CTCAATCAGAAAATTTGCTTGTGCTTCTGGATGCAACAAATTTATGAGCTATTCAA 1680
Db 1621 CTCAATCAGAAAATTTGCTTGTGCTTCTGGATGCAACAAATTTATGAGCTATTCAA 1680
Qy 1681 GTGACTTTAGGCAGATCTTTTCCAGCATGCCCTTAGAGAGAACCTGACTCGGAGATC 1740
Db 1681 GTGACTTTAGGCAGATCTTTTCCAGCATGCCCTTAGAGAGAACCTGACTCGGAGATC 1740
Qy 1741 CCTGCTCATAAAGCAGGTTAATGCTTTGGGATGCAAGCTTTATAAAGTCCCGTAAGA 1800
Db 1741 CCTGCTCATAAAGCAGGTTAATGCTTTGGGATGCAAGCTTTATAAAGTCCCGTAAGA 1800
Qy 1801 TCCCTCATGAATTTGATCTTCAGCAGTTTATTTATGAGAGTGAATACAAAAAACTCTTTG 1860
Db 1801 TCCCTCATGAATTTGATCTTCAGCAGTTTATTTATGAGAGTGAATACAAAAAACTCTTTG 1860
Qy 1861 CTATGGAATTTGCTGAAGTATTAACAACTGCGAGAAAGATATATCAGTGATCATG 1920
Db 1861 CTATGGAATTTGCTGAAGTATTAACAACTGCGAGAAAGATATATCAGTGATCATG 1920
Qy 1921 ACAGAGTATCTCTATTAACCTGCTTTCAGTTTCAGATGTTTACTGTTCTTACTCTGGCTC 1980
Db 1921 ACAGAGTATCTCTATTAACCTGCTTTCAGTTTCAGATGTTTACTGTTCTTACTCTGGCTC 1980
Qy 1981 GACATCTTTATGGAAGCAGAAATGTTATCTCTGTCATTACTGAAACTCTGCTAGAGTTT 2040
Db 1981 GACATCTTTATGGAAGCAGAAATGTTATCTCTGTCATTACTGAAACTCTGCTAGAGTTT 2040
Qy 2041 TACCTGAGTACTTGGACAGGAACTAATAATTCACCTTCCAGGGTTATAGCCAGGACAAAT 2100
Db 2041 TACCTGAGTACTTGGACAGGAACTAATAATTCACCTTCCAGGGTTATAGCCAGGACAAAT 2100
Qy 2101 TGGCAAGAGTATATGCACTTAAAGTATATCTGATCAGCAAAACCCACAA 2160
Db 2101 TGGCAAGAGTATATGCACTTAAAGTATATCTGATCAGCAAAACCCACAA 2160
Qy 2161 TATGGACAGAAAGTAAAGATGCAAGTCTTGAAGGGTTTCGATCTTTTGAAGATTC 2220
Db 2161 TATGGACAGAAAGTAAAGATGCAAGTCTTGAAGGGTTTCGATCTTTTGAAGATTC 2220
Qy 2221 TTACTGTTATGCAAGGAAATGGAAGAAATCCGAGACAGGTTGGGCAACACATTGAAGTGG 2280
Db 2221 TTACTGTTATGCAAGGAAATGGAAGAAATCCGAGACAGGTTGGGCAACACATTGAAGTGG 2280
Qy 2281 ATCTCTGATTTGGGAGCTGCCATTCCTATACAGATGCAATTCGAAGAAATTTTACTCATGT 2340
Db 2281 ATCTCTGATTTGGGAGCTGCCATTCCTATACAGATGCAATTCGAAGAAATTTTACTCATGT 2340
Qy 2341 TCCAAAGAGTGGTGTGCTTGTGATGAAGAACTCTTACTTGTGGCTTATAAAGAAATGTCA 2400
Db 2341 TCCAAAGAGTGGTGTGCTTGTGATGAAGAACTCTTACTTGTGGCTTATAAAGAAATGTCA 2400
Qy 2401 AAGCTGTGATGAGGTGCAAGTACCAAGTTCATATCTAGTAGCAACAGTAGTACATCGT 2460
Db 2401 AAGCTGTGATGAGGTGCAAGTACCAAGTTCATATCTAGTAGCAACAGTAGTACATCGT 2460

QY 2461 GTGGACATAGTTTGGAAACAAAGTCCTACAGATATCTGAGGATCTTGTAAAGCATACATC 2520
Db 2461 GTGGACATAGTTTGGAAACAAAGTCCTACAGATATCTGAGGATCTTGTAAAGCATACATC 2520
QY 2521 TGCCACTCTCTAGGACCCCTGCTGCTCTTCTATGACGTTTAAAGCAGGCTGGGTGCTGTTT 2580
Db 2521 TGCCACTCTCTAGGACCCCTGCTGCTCTTCTATGACGTTTAAAGCAGGCTGGGTGCTGTTT 2580
QY 2581 CAAGACTGCATGAATTTGTGCTCTTGTAGGACTTTTCAAGTAGAGTACTAGTGGAAATATC 2640
Db 2581 CAAGACTGCATGAATTTGTGCTCTTGTAGGACTTTTCAAGTAGAGTACTAGTGGAAATATC 2640
QY 2641 CTTTACGTTGCTGGTGTGGTGGTCCAGGTTGCTGAGATGTCGGGAGAAATGAC 2700
Db 2641 CTTTACGTTGCTGGTGTGGTGGTCCAGGTTGCTGAGATGTCGGGAGAAATGAC 2700
QY 2701 TGTCTCTTATTAGCCAGGTTTATTAACCAAGATGTTAAAGTGCAGAGAAATGATG 2760
Db 2701 TGTCTCTTATTAGCCAGGTTTATTAACCAAGATGTTAAAGTGCAGAGAAATGATG 2760
QY 2761 ATAAAGATPATCATGCTTCAGATGCTGATCTTTAATGGATCCCAATAGTCTTGT 2820
Db 2761 ATAAAGATPATCATGCTTCAGATGCTGATCTTTAATGGATCCCAATAGTCTTGT 2820
QY 2821 TACTGGTACTTCAGAGGTATGAACTTGGCGAGGCTTTTAAACAGACCATATCTACAAAG 2880
Db 2821 TACTGGTACTTCAGAGGTATGAACTTGGCGAGGCTTTTAAACAGACCATATCTACAAAG 2880
QY 2881 ACCAGGATTTGATTAACCAATATATACATTAATAGAGAAATGCTTCAGGTCCTCATCT 2940
Db 2881 ACCAGGATTTGATTAACCAATATATACATTAATAGAGAAATGCTTCAGGTCCTCATCT 2940
QY 2941 ATATTGTTGGGTGAGGCTTATGCTGAGGTGGGAAATGTGACCAAGAGAGGTCAAA 3000
Db 2941 ATATTGTTGGGTGAGGCTTATGCTGAGGTGGGAAATGTGACCAAGAGAGGTCAAA 3000
QY 3001 TGAGAGAAATCATTCATCTGCTTTGCAATGCAATGCCATGCCACAGTGCCATTTGCCAAA 3060
Db 3001 TGAGAGAAATCATTCATCTGCTTTGCAATGCAATGCCATGCCACAGTGCCATTTGCCAAA 3060
QY 3061 ATTTACCTTGAGATGAATAATTAAGAACTGCTGAGAAATGTCAATAACAAAGTGGCCA 3120
Db 3061 ATTTACCTTGAGATGAATAATTAAGAACTGCTGAGAAATGTCAATAACAAAGTGGCCA 3120
QY 3121 CATTTAAGAAACCAAGGTATCAGGCCATGGAGTTTATGAATGAATGAATCACTGA 3180
Db 3121 CATTTAAGAAACCAAGGTATCAGGCCATGGAGTTTATGAATGAATGAATCACTGA 3180
QY 3181 AAGACTTCAATATGATCTTTTATCAATTAATCAAAACCCAGCATAGCAAGGCTGAACATA 3240
Db 3181 AAGACTTCAATATGATCTTTTATCAATTAATCAAAACCCAGCATAGCAAGGCTGAACATA 3240
QY 3241 TGCAAGAAAGGAGAAACCAAGAAACCAAGATGAAGCATTCGCCACCAACCCCTC 3300
Db 3241 TGCAAGAAAGGAGAAACCAAGAAACCAAGATGAAGCATTCGCCACCAACCCCTC 3300
QY 3301 CTGAATTCGCTGCTCTTTCAGCAAGATGATTAACCTTCTCACTGCTGATATCATGATG 3360
Db 3301 CTGAATTCGCTGCTCTTTCAGCAAGATGATTAACCTTCTCACTGCTGATATCATGATG 3360
QY 3361 ACATTTCTCAGGACCGTATTTGAGCGGGCAATAGACACAGATTTCTAACTTTGGACCGAAG 3420
Db 3361 ACATTTCTCAGGACCGTATTTGAGCGGGCAATAGACACAGATTTCTAACTTTGGACCGAAG 3420
QY 3421 GGATGCTCCAAATGGCTTTTATTTCTGGCATTTGGGTTTACTAGAGAGAGCAACAGC 3480
Db 3421 GGATGCTCCAAATGGCTTTTATTTCTGGCATTTGGGTTTACTAGAGAGAGCAACAGC 3480
QY 3481 TTCAAAAGCTCTTGAAGAGAGTAAATTTGATCTTTTATCATTAAGGCTTCAAGATTGG 3540
Db 3481 TTCAAAAGCTCTTGAAGAGAGTAAATTTGATCTTTTATCATTAAGGCTTCAAGATTGG 3540
QY 3541 GAACTTCAGCCATGAATATACAAATGCTTTTGGAAAACTCAAGGAATTTCCCGCTTAG 3600

Db 3541 GAACTTCAGCCATGAATATACAAATGCTTTTGGAAAACTCAAGGAATTTCCCGCTTAG 3600
QY 3601 AAGCCAGAGAGCATGATTAACGTTGGATATCTTCAGATGTTTGGACACAGTGAAGCGATTAA 3660
Db 3601 AAGCCAGAGAGCATGATTAACGTTGGATATCTTCAGATGTTTGGACACAGTGAAGCGATTAA 3660
QY 3661 GAGAAAAATCTTGTGTTTAAATTTGTAGCAACCATCAGGATCGGAATCTTATTAAAGATGATG 3720
Db 3661 GAGAAAAATCTTGTGTTTAAATTTGTAGCAACCATCAGGATCGGAATCTTATTAAAGATGATG 3720
QY 3721 AGATTACTCATGATTAAGAAAAAGCAGAAAGAAAAAGAAAGCTGAAAGCTCTAGGCTTAC 3780
Db 3721 AGATTACTCATGATTAAGAAAAAGCAGAAAGAAAAAGAAAGCTGAAAGCTCTAGGCTTAC 3780
QY 3781 ATGCCCAGAGAGATCATGGCTCAGATGCTGCTTTACAGAAAAAATCTTCAATGAAACTCAT 3840
Db 3781 ATGCCCAGAGAGATCATGGCTCAGATGCTGCTTTACAGAAAAAATCTTCAATGAAACTCAT 3840
QY 3841 AACTCATGATGACATATACATCAGAAATGCTTGGGAAAGAGATTCATTTATGGAGGAAAG 3900
Db 3841 AACTCATGATGACATATACATCAGAAATGCTTGGGAAAGAGATTCATTTATGGAGGAAAG 3900
QY 3901 AGAGCACCCAGCAGTCAAGTACTTCTAGAAATGCTTGGGCTCTTAAACGGGCTCCAT 3960
Db 3901 AGAGCACCCAGCAGTCAAGTACTTCTAGAAATGCTTGGGCTCTTAAACGGGCTCCAT 3960
QY 3961 CTGTTACTGAAAAAGGAGGCTGCTGACGTCATCTTTTGCAGAAAGAAACAGAGAGTGAATA 4020
Db 3961 CTGTTACTGAAAAAGGAGGCTGCTGACGTCATCTTTTGCAGAAAGAAACAGAGAGTGAATA 4020
QY 4021 TAGAAAAATATGCAATGCTTATCGGCTGCTGTCAGAAATCTTACGCTTAAACCCAGC 4080
Db 4021 TAGAAAAATATGCAATGCTTATCGGCTGCTGTCAGAAATCTTACGCTTAAACCCAGC 4080
QY 4081 ACAGGGAAACCCATAGAACTCTCAGAGAGAGCCCTAGAGCCACTTTTCATGATGATCCAG 4140
Db 4081 ACAGGGAAACCCATAGAACTCTCAGAGAGAGCCCTAGAGCCACTTTTCATGATGATCCAG 4140
QY 4141 ACTTGGCATATGGAACCTTATACAGAAAGCTGCTGCTGATGTAATGCAAGCAGTGTCTGGC 4200
Db 4141 ACTTGGCATATGGAACCTTATACAGAAAGCTGCTGCTGATGTAATGCAAGCAGTGTCTGGC 4200
QY 4201 AGAAGTATTTTGAAGCTGTACAGCTGCTCAGCAGCGCATTCATGTTGACCTTTTGTG 4260
Db 4201 AGAAGTATTTTGAAGCTGTACAGCTGCTCAGCAGCGCATTCATGTTGACCTTTTGTG 4260
QY 4261 ACTTGGAAAGTGGAGAAATATCTTTGCTCTTTGCAATCTCTGTCGAATCTGTCATCTGATCC 4320
Db 4261 ACTTGGAAAGTGGAGAAATATCTTTGCTCTTTGCAATCTCTGTCGAATCTGTCATCTGATCC 4320
QY 4321 CCATTTATTTCTTTCGCAACCTCAAAAGATTAACAGTGAAGATGTCGATGCTCTTGTCAAC 4380
Db 4321 CCATTTATTTCTTTCGCAACCTCAAAAGATTAACAGTGAAGATGTCGATGCTCTTGTCAAC 4380
QY 4381 TTTTGGACCTGGCAGCGTGGATACAGATGTTTGGCCAGAAATTCAGGTTTATTAATATA 4440
Db 4381 TTTTGGACCTGGCAGCGTGGATACAGATGTTTGGCCAGAAATTCAGGTTTATTAATATA 4440
QY 4441 GACATGCTTAAAGGAGAAAAACCAATTCCTATTTTCTTTAATCAAGGAATGGGAGATTCTA 4500
Db 4441 GACATGCTTAAAGGAGAAAAACCAATTCCTATTTTCTTTAATCAAGGAATGGGAGATTCTA 4500
QY 4501 CTTTGGAGTTTCCATTCCTGAGTTTGGGCTTGGATCTTCGATTAATAATATTCAATA 4560
Db 4501 CTTTGGAGTTTCCATTCCTGAGTTTGGGCTTGGATCTTCGATTAATAATATTCAATA 4560
QY 4561 GCATCAAGGAAATGTTTATCTCTTTGCAACAAATTTATAGAAATTTGGAATGGAAGTGC 4620
Db 4561 GCATCAAGGAAATGTTTATCTCTTTGCAACAAATTTATAGAAATTTGGAATGGAAGTGC 4620
QY 4621 CACTGATGAAGGGATCTCTGAGTCCCATGCTGACCTGGAGCACCTGCGCTTCACTA 4680

Db 4621 CACCTGATGAAGGATCCTCGAGTCCCACTGCTGACCTGAGCAGCACCTGCGCTTTCACTA 4680
Qy TCCAGGCAATTTGAAATCTATTTGGAGATGAAGAAACCTCTCTTTGGAGCACTTCAAA 4740
Db TCCAGGCAATTTGAAATCTATTTGGAGATGAAGAAACCTCTCTTTGGAGCACTTCAAA 4740
Qy ATAGCAGCAATAATGGTCTGAAGCAATTAATGCAAGTTCGAGTTGCAAGAGGATTAACCT 4800
Db ATAGCAGCAATAATGGTCTGAAGCAATTAATGCAAGTTCGAGTTGCAAGAGGATTAACCT 4800
Qy GTCTCAGGTCCTGATACAGAAACATCTGGTTCGTTCTTATCAATGTTCTTCTCTAACA 4860
Db GTCTCAGGTCCTGATACAGAAACATCTGGTTCGTTCTTATCAATGTTCTTCTCTAACA 4860
Qy TAAATCAGAGATACACCATGCTTCTGCTATAGATCTGTTCAATGTTTGGTGGTG 4920
Db TAAATCAGAGATACACCATGCTTCTGCTATAGATCTGTTCAATGTTTGGTGGTG 4920
Qy CTGTGTTAGCAATCCCATCTTGTATTTGGGATGACCTGTTGATCTGAGGCTTCTTCAG 4980
Db CTGTGTTAGCAATCCCATCTTGTATTTGGGATGACCTGTTGATCTGAGGCTTCTTCAG 4980
Qy TTAGTCTTCTTAACACCATTTATCTTCCATTTGATCAACATGAGCAACATGCTTC 5040
Db TTAGTCTTCTTAACACCATTTATCTTCTTCCATTTGATCAACATGAGCAACATGCTTC 5040
Qy AGATACACTTACAGTAGACACAGGCTTACCCCTGCTCAGGTTTCAGAAACAGTAGAAG 5100
Db AGATACACTTACAGTAGACACAGGCTTACCCCTGCTCAGGTTTCAGAAACAGTAGAAG 5100
Qy AGGCTCATTTCCGCACTCTTCTTTGAGAAATTTCTCAATATACAGTAGGCTTCCATTTG 5160
Db AGGCTCATTTCCGCACTCTTCTTTGAGAAATTTCTCAATATACAGTAGGCTTCCATTTG 5160
Qy GGTGTGATTAATCTGGCTGATTTGTTGGGCTCTCACTGAAGAAATGGCAATCCCTTATC 5220
Db GGTGTGATTAATCTGGCTGATTTGTTGGGCTCTCACTGAAGAAATGGCAATCCCTTATC 5220
Qy TTGCTGTGCTGCAATGTTTTCATTTTCTTGGGTAACCTCGGCTGAGCACTGC 5280
Db TTGCTGTGCTGCAATGTTTTCATTTTCTTGGGTAACCTCGGCTGAGCACTGC 5280
Qy ATACCAATTTCTGAGAGGAGTACAGTCACTCTGTAGTATCTATCTTTTACTCTACAA 5340
Db ATACCAATTTCTGAGAGGAGTACAGTCACTCTGTAGTATCTATCTTTTACTCTACAA 5340
Qy ATTTGTTCTCTCTCTTCCAGGAATATTTGGGTAACCTCGGCTGAGCACTGC 5400
Db ATTTGTTCTCTCTCTTCCAGGAATATTTGGGTAACCTCGGCTGAGCACTGC 5400
Qy GTGCAATCTCTGCAATGTTTGAAGCAAAACACCGTGTGAGTACCTTA 5460
Db GTGCAATCTCTGCAATGTTTGAAGCAAAACACCGTGTGAGTACCTTA 5460
Qy GAAAGAAATAGTTTGTAGAGCTTCTGATGACTATAGTCTGCTCTGATCAAGCTT 5520
Db GAAAGAAATAGTTTGTAGAGCTTCTGATGACTATAGTCTGCTCTGATCAAGCTT 5520
Qy CTCAATTCAGGTGCCACGGTCTGAGATGATGAGCAAGCATCTGCTCTGCTTC 5580
Db CTCAATTCAGGTGCCACGGTCTGAGATGATGAGCAAGCATCTGCTCTGCTTC 5580
Qy TCTGTGGGGCTACTATGTTCTCAGACATTTGCTGCGAGAAATTTGTAACGGGAG 5640
Db TCTGTGGGGCTACTATGTTCTCAGACATTTGCTGCGAGAAATTTGTAACGGGAG 5640
Qy AGGTTGAGCTTGCATTTTTCAGCACTTCTGCTGAGCGGAGTCTGCAATTTCTTAA 5700
Db AGGTTGAGCTTGCATTTTTCAGCACTTCTGCTGAGCGGAGTCTGCAATTTCTTAA 5700
Qy AAATCAGAGATGCGGAGTGTCTGTTGAAGGTAAGCCAGAGGCTGTGCTTATCAG 5760
Db AAATCAGAGATGCGGAGTGTCTGTTGAAGGTAAGCCAGAGGCTGTGCTTATCAG 5760

Qy 5761 CTCCTACTTTGGATGAATATGGAGAAACAGACCCCTGGCTTGAAGAGGGGCAACCCCTTC 5820
Db CTCCTACTTTGGATGAATATGGAGAAACAGACCCCTGGCTTGAAGAGGGGCAACCCCTTC 5820
Qy 5821 ATTTATCTCTGAGCGGTATCGGAAGCTCCATTTGGTCTGGCAACAACACTGCAATTATAG 5880
Db ATTTATCTCTGAGCGGTATCGGAAGCTCCATTTGGTCTGGCAACAACACTGCAATTATAG 5880
Qy 5881 AAGGATTTCTAGGAGCAAGAGACTAATCAGATGTTTATTTGGATTCAACTGGCAGTTAC 5940
Db AAGGATTTCTAGGAGCAAGAGACTAATCAGATGTTTATTTGGATTCAACTGGCAGTTAC 5940
Qy 5941 TGTGAGCTTCCAACTCTCGCTCAAGACAATCAAAATGACACAGTGTAAAGGCTGATTC 6000
Db TGTGAGCTTCCAACTCTCGCTCAAGACAATCAAAATGACACAGTGTAAAGGCTGATTC 6000
Qy 6001 AAAATTATGAAATCTTTCTGAGGGCTGGGAAAGTATTTGGAGGCTCTTTGCTCCATGTC 6060
Db AAAATTATGAAATCTTTCTGAGGGCTGGGAAAGTATTTGGAGGCTCTTTGCTCCATGTC 6060
Qy 6061 CAGGTTCACTTACATCAATTAATAATTTCTTAATGGAGTATTTGCTTCAATTAGCAACA 6120
Db CAGGTTCACTTACATCAATTAATAATTTCTTAATGGAGTATTTGCTTCAATTAGCAACA 6120
Qy 6121 TATGCTTTCAGAGAAAGGACATAGATCAATCTGTTTATGCTAGTAGTATTTCCAGGA 6180
Db TATGCTTTCAGAGAAAGGACATAGATCAATCTGTTTATGCTAGTAGTATTTCCAGGA 6180
Qy 6181 ATTTATTTCCCTTCAATTTGCTCAATTTTCAATTTTATTTCACTCACTTGGTAGAGAA 6240
Db ATTTATTTCCCTTCAATTTGCTCAATTTTCAATTTTATTTTCACTCACTTGGTAGAGAA 6240
Qy 6241 TCAGTCAAAAGTGTAGACATTTATGTTGGTTAACTCTTCTGCAATTTTGTATTT 6300
Db TCAGTCAAAAGTGTAGACATTTATGTTGGTTAACTCTTCTGCAATTTTGTATTT 6300
Qy 6301 GGTGTTTT 6308
Db GGTGTTTT 6308

RESULT 2
AV061886 5250 bp mRNA linear PRI 30-DEC-2002
LOCUS Homo sapiens ubiquitin ligase E3 alpha-I mRNA, complete cds.
DEFINITION
ACCESSION AY061886
VERSION AY061886.1 GI:27434483
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 5250)
AUTHORS Han,H.Q. and Kwak,K.
TITLE Full-length human ubiquitin ligase E3 alpha-I (E3 alpha)
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 5250)
AUTHORS Han,H.Q. and Kwak,K.
TITLE Direct Submission
JOURNAL Submitted (05-NOV-2001) Research, Amgen Inc., One Amgen Center
Drive, Thousand Oaks, CA 91320, USA

FEATURES
source
1. .5250
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1. .5250
/codon_start=1
/product="ubiquitin ligase E3 alpha-I"
/protein_id="AAU32103.1"
/db_xref="GI:27434484"
/translation="MADEAGGTERMEISABELPQTPORLASWWDQVDFYTFALHILA

OLVPEIYFAEMDPDLKOEBSVOMSIPTLEWYLFGEEDPDICLEKLKHSQAPQLCGRV
FKSGETTSYSCRDCAIDPTCVLCMDQFQSVHKNHRYKQHTSTGGGFCDCGDEAWKQ
PFCVNHPEGRAGCTIKENSRPLNEEVIPOARKIPSPVYKYVVMETIWEKEELPPELQ
IRKRENYCYVLFNDDEHSHYHVIYSLOALDCELAELQHTTADKEGRRAVAKAG
AACQBAKEDIKSHSNVSOHPHLHVEVLHSEIMAHQKFAIRLGSWMNKIYSYSDPQOI
FCOACLRBPDSQCLISRLMLWDAKYKGAARKILHLEIPSPFPMEMBYKKLPAMFB
VKYKLOKEIYISDDHDSRISITALSVMQFTPTLARIHLEBQNVISVITETLLEVL
EVLDRNKNFNGQYSQDKLGRVAVICDLKYILLISKPTIWIETRLRMQPLEGPRSPKI
LTCMOGBEIRKQVQGHLEVDPDWEAAITAIQMLKNILLMPOBWCADBEILLVAIKE
CHKAMRSTSEKQIPESSKTVOQSGHSLTYSRVSEDLVSHLPSRSLAGLHVRSL
LGAVERLHVEFQVEVLVEPLRCLVLAQVVAEMWRNGLSLISQVFFYQDVK
CREENMYDKDILMLOICASLMDPNKLLVLORYELAEAFNKITSKDOODLKOYNVTL
BEMLOVLIVYIGERVVPGVGNVKEBWTREIHLICIEPMHPSAIAKNLEPENNET
GLEWINKVAFTKPGVSGHGYELKDSLDKDPNMYFYHYSKTQHSKASHMQKRRQ
ENKDBALPPPPPEFCAPFSKVINLCIDIMYIILKTFERAILDSDSNLWTEGLQMA
PHILALGULLEKQLOKQPEEVEITDFYHKASRLGSSAMNTOMLLEKLGIPQLEGO
DMITWILQFDTVTKLRKSCCLIVATTSGBSISIKNDEITHDKEKAEKRRKAEARLHR
OKIMQMSALQNFETETHKLYMDNTSEMPGKEDSIMEBESTPAVSDYSRIALGPKRG
SVTEKEVLTCILCOREQVKIENNAWLSACVOKSTALTTOHREKPIELSGEALDPLM
DPDLAYGTTSGCHVMNAVWCKYFPEAVQLSSQORIHVDLPDLSEGYLCPKSKSLC
NTVPIIPLQPKINSENADALQILLTARWITQTLARIISGNIHRHAKGENPIPIFFN
QGMGSTLEFHSILSPGVESSIKYSNSIKEMVILFATTIYRIGLVPDPDRPVPLM
TSCACFTAQIENLQEGKPLFGALQNRHNGLKALMQPAVQORIICPOVLIOKHL
VLLSVLPNLSKEDTCLLSIDLPHLVGAVLAPPSLVDDPDVQLQSPSSVSNHL
WTLHLITWAMLIQILLTVDVTCPLAQVQEDSEAHSSAPFAEISQYTSIGICDIPG
WYLSWLSKNGITPILQCAALFFHYLLGVTPPELHNSABGEYSALCSYLSUPTNLF
LFQBYWDTVPLQORCADPALLNCLKQKNTVYPRKRNLSLIEPLDPDYSCLLNQASH
FRCPASADRRKHPVLICFPGAILCSQNICQBIVNGBEVGACITPHALHCGAGVICPL
KIRCVVLVBGARGCAYPAPYLDEYGETDPLGRGNPLHLRERYRKLHLVMOQHC
IIEEFARSOETNOMLFGFNWOLL"

ORIGIN

Query Match		83.2%;	Score 5250;	DB 9;	Length 5250;
Best Local Similarity		100.0%;	Pred. No. 0;		
Matches 5250;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	696	ATGGCGGACGAGGAGCTGGAGTACTGAGGATGGAATCAGCGCGGAGTTACCCAG	755		
Db	1	ATGGCGGACGAGGAGCTGGAGTACTGAGGATGGAATCAGCGCGGAGTTACCCAG	60		
QY	756	ACCCCTCAGCGTCTGGCATCTTGTGGGATCAGCAAGTTGATTTTATCTGCTTTCTTG	815		
Db	61	ACCCCTCAGCGTCTGGCATCTTGTGGGATCAGCAAGTTGATTTTATCTGCTTTCTTG	120		
QY	816	CATCATTTGGCAAAATGGTCCAGAAATTTACTTGTGAAATGGACCCAGACTTGGAA	875		
Db	121	CATCATTTGGCAAAATGGTCCAGAAATTTACTTGTGAAATGGACCCAGACTTGGAA	180		
QY	876	AAGCAGGAGGAAGGTACAAATGTCAATATTCATCTCCACTGGATGTACTTATTTGA	935		
Db	181	AAGCAGGAGGAAGGTACAAATGTCAATATTCATCTCCACTGGATGTACTTATTTGA	240		
QY	936	GAAGATCCAGATATTGTCTTACAGAAATTTGAGACAGTGGAGCATTTTCAGCTTTGTGG	995		
Db	241	GAAGATCCAGATATTGTCTTACAGAAATTTGAGACAGTGGAGCATTTTCAGCTTTGTGG	300		
QY	996	AGGGTTTTCAAAGTGGAGAGCAACCTATTCTTTCAGGGAATGTGCAATTTGATCCAACA	1055		
Db	301	AGGGTTTTCAAAGTGGAGAGCAACCTATTCTTTCAGGGAATGTGCAATTTGATCCAACA	360		
QY	1056	TGTGTACTCTGTATGGACTGCTTCAGGACAGTGTTCATAAATAATCATCGTTACAAGATG	1115		
Db	361	TGTGTACTCTGTATGGACTGCTTCAGGACAGTGTTCATAAATAATCATCGTTACAAGATG	420		
QY	1116	CATACCTTCTACTGGAGGAGGTCTGTGATCTGTGGAGACACAGAGGCATGGAAACTGGC	1175		
Db	421	CATACCTTCTACTGGAGGAGGTCTGTGATCTGTGGAGACACAGAGGCATGGAAACTGGC	480		
QY	1176	CCTTTTGTGTAAATCATGAACCTCGAAGAGCAGGTACTATATAAGAGAAATTCAGCTGT	1235		
Db	481	CCTTTTGTGTAAATCATGAACCTCGAAGAGCAGGTACTATATAAGAGAAATTCAGCTGT	540		
QY	1236	CCGTTGAATGAAGGTAATTTGTCCAGCCAGGAAAAATATTTCCTTCAGTGATAAAAATAT	1295		

Db	541	CCGTTGAATGAAGGTAATTTGTCCAGCCAGGAAAAATATTTCCTTCAGTGATAAAAATAT	600		
QY	1296	GTCTAGTAAATGACTATATGGGAAGAGGAAAAAGAACTCCCTCTCGAACTCCAGATAAGG	1355		
Db	601	GTCTAGTAAATGACTATATGGGAAGAGGAAAAAGAACTCCCTCTCGAACTCCAGATAAGG	660		
QY	1356	GAGAAAAATGAAGATACATATTGTCTCTTTTCAATGATGAACACCATTCATATGACCAAC	1415		
Db	661	GAGAAAAATGAAGATACATATTGTCTCTTTTCAATGATGAACACCATTCATATGACCAAC	720		
QY	1416	GTCTATATACAGCTCAAAAGAGCTCTTGACTGTGTAGCTCGCAGAGGCCAGTTGCATACC	1475		
Db	721	GTCTATATACAGCTCAAAAGAGCTCTTGACTGTGTAGCTCGCAGAGGCCAGTTGCATACC	780		
QY	1476	ACTGCCATTGACAAAGAGGGTGTGCGGGCTGTGTTAAAGCGGGAGCTTATGCTGCTTCCAC	1535		
Db	781	ACTGCCATTGACAAAGAGGGTGTGCGGGCTGTGTTAAAGCGGGAGCTTATGCTGCTTCCAC	840		
QY	1536	GAGCAAAAGGAGATATATAGAGTCATTCAGAAAAATGTCTCTCAACATCCACTTCATGTA	1595		
Db	841	GAGCAAAAGGAGATATATAGAGTCATTCAGAAAAATGTCTCTCAACATCCACTTCATGTA	900		
QY	1596	GAGTATTACACTCAGAGATTATGCTCATCAGAAAAATTTGCTTTCGCTCTGCTTCTCTGG	1655		
Db	901	GAGTATTACACTCAGAGATTATGCTCATCAGAAAAATTTGCTTTCGCTCTGCTTCTCTGG	960		
QY	1656	ATGAACAAAAATTTATGAGCTATTCAAGTGACTTTTAGGCGAGATCTTTTGGCAAGCATGCC	1715		
Db	961	ATGAACAAAAATTTATGAGCTATTCAAGTGACTTTTAGGCGAGATCTTTTGGCAAGCATGCC	1020		
QY	1716	AGAGAAAGAACTGACTCGGAGAAATCCCTGTCTCATAGCAGGTTAATGCTTTGGGATGCA	1775		
Db	1021	AGAGAAAGAACTGACTCGGAGAAATCCCTGTCTCATAGCAGGTTAATGCTTTGGGATGCA	1080		
QY	1776	AAGCTTTATAAGGTCGCCGTAAGATCCTTCATGAATTTGATCTTCAGCAGTTTTTTTATG	1835		
Db	1081	AAGCTTTATAAGGTCGCCGTAAGATCCTTCATGAATTTGATCTTCAGCAGTTTTTTTATG	1140		
QY	1836	GAGATGGAATACAAAAAACTCTTTGCTATGGAATTTGTGAAAGTATTATAAACAATGCGAG	1895		
Db	1141	GAGATGGAATACAAAAAACTCTTTGCTATGGAATTTGTGAAAGTATTATAAACAATGCGAG	1200		
QY	1896	AAAGATATATCAGTGATGATCATGACAGAGTATCTCTATAAATGCGACTTTTCAGTTGAG	1955		
Db	1201	AAAGATATATCAGTGATGATCATGACAGAGTATCTCTATAAATGCGACTTTTCAGTTGAG	1260		
QY	1956	ATGTTTACTGTTCTTCTCTGCTCGACATCTTATTGGAAGCAGAAATGTTATCTCTGTC	2015		
Db	1261	ATGTTTACTGTTCTTCTCTGCTCGACATCTTATTGGAAGCAGAAATGTTATCTCTGTC	1320		
QY	2016	ATTACTGAAACTCTGCTAGAAAGTTTACTCTGAGTACTTTGGGACAGGAAATATAAATTTCAAC	2075		
Db	1321	ATTACTGAAACTCTGCTAGAAAGTTTACTCTGAGTACTTTGGGACAGGAAATATAAATTTCAAC	1380		
QY	2076	TTCCAGGGTTATAGCCAGGACAAATTTGGGAAGAGTATATGCAATATGAGCACTTAAG	2135		
Db	1381	TTCCAGGGTTATAGCCAGGACAAATTTGGGAAGAGTATATGCAATATGAGCACTTAAG	1440		
QY	2136	TATATCTGTATCAGCAAAACCAATATGGAACAGAAATTAAGATGCAAGTTCTCTGAA	2195		
Db	1441	TATATCTGTATCAGCAAAACCAATATGGAACAGAAATTAAGATGCAAGTTCTCTGAA	1500		
QY	2196	GGTTTTCAGATCTTTTGAAGATTTCTTACCTGTATGCGAGGAAATGGAAGAAATCCGAAGA	2255		
Db	1501	GGTTTTCAGATCTTTTGAAGATTTCTTACCTGTATGCGAGGAAATGGAAGAAATCCGAAGA	1560		
QY	2256	CAGGTGGGCAACACATTTGAAGTGGATCTCTGATTTGGGAGGCTGCCATTTGCTATACAGATG	2315		
Db	1561	CAGGTGGGCAACACATTTGAAGTGGATCTCTGATTTGGGAGGCTGCCATTTGCTATACAGATG	1620		
QY	2316	CAATTTGAAGATATTTTACTCATGTTCCAGAGTGGTGTGCTTGTGATGAAGAACTCTTA	2375		

Db 1621 CAATTTGAAGAAATATTTTACTCATGTTTCAAGAGTGGTGTGCTGTGATGAAGAACTCTTA 1680
Qy 2376 CTTGTGGCTTATAAAGAAATGTCAAAAGCTGTGATGAGGTGAGTACAGATTTTCAATATCT 2435
Db 1681 CTTGTGGCTTATAAAGAAATGTCAAAAGCTGTGATGAGGTGAGTACAGATTTTCAATATCT 1740
Qy 2436 AGTAGCAGACAGTAGTAGTAACAATCGTGTGACATAGTTTGGAAAAAACAAGTCTCTACAGAGTA 2495
Db 1741 AGTAGCAGACAGTAGTAGTAACAATCGTGTGACATAGTTTGGAAAAAACAAGTCTCTACAGAGTA 1800
Qy 2496 TCTGAGGATCTTTGTAAGCATACATCTGCCACTCTCTAGGACCCCTTGCTGGTCTTTCATGTA 2555
Db 1801 TCTGAGGATCTTTGTAAGCATACATCTGCCACTCTCTAGGACCCCTTGCTGGTCTTTCATGTA 1860
Qy 2556 CGTTTAAAGCAGGCTGGGTGCTGTTTCAAGACCTGCATGAATTTGTGTCTTTTGGAGACTTTT 2615
Db 1861 CGTTTAAAGCAGGCTGGGTGCTGTTTCAAGACCTGCATGAATTTGTGTCTTTTGGAGACTTTT 1920
Qy 2616 CAAGTAGAGGTACTAGTGGAAATATCCTTTTACGTTGCTCTGGTGTGGTTGGCCAGGTGCTT 2675
Db 1921 CAAGTAGAGGTACTAGTGGAAATATCCTTTTACGTTGCTCTGGTGTGGTTGGCCAGGTGCTT 1980
Qy 2676 GCTGAGATGTGGCGAAGAAATGGACTGTCTTTATTTAGCCAGGTGTTTTTATTTACCAAGAT 2735
Db 1981 GCTGAGATGTGGCGAAGAAATGGACTGTCTTTATTTAGCCAGGTGTTTTTATTTACCAAGAT 2040
Qy 2736 GTTAAGTGCAGAGAAAGAAATGTATGAATAAGATATCATCTGCTTCAGATTTGGTGCATCT 2795
Db 2041 GTTAAGTGCAGAGAAAGAAATGTATGAATAAGATATCATCATGCTTTCAGATTTGGTGCATCT 2100
Qy 2796 TTAATGGATCCCAATAAGTCTTGTGTTACTGTGTACTTCAGAGGTATGAACCTTGGCCAGGCT 2855
Db 2101 TTAATGGATCCCAATAAGTCTTGTGTTACTGTGTACTTCAGAGGTATGAACCTTGGCCAGGCT 2160
Qy 2856 TTTAAACAAGACCATCTACAAAAGACAGGATTTGATTTAAACAATAATAATACATAATA 2915
Db 2161 TTTAAACAAGACCATCTACAAAAGACAGGATTTGATTTAAACAATAATAATACATAATA 2220
Qy 2916 GAAGAAATGCTTCAGGCTCTCATATATATTTGGGTGAGCGTTATGTACCTGGAGTGGGA 2975
Db 2221 GAAGAAATGCTTCAGGCTCTCATATATATTTGGGTGAGCGTTATGTACCTGGAGTGGGA 2280
Qy 2976 AATGTGACCAAGAGAGGTCAATGAGAGAAATCAATTCATCTGCTTTGCAATGAACCC 3035
Db 2281 AATGTGACCAAGAGAGGTCAATGAGAGAAATCAATTCATCTGCTTTGCAATGAACCC 2340
Qy 3036 ATGCCACACAGTCCCAATTTGCCAAAATTTTACCTGAGAAATGAATAATAATGAACCTGGCTTA 3095
Db 2341 ATGCCACACAGTCCCAATTTGCCAAAATTTTACCTGAGAAATGAATAATAATGAACCTGGCTTA 2400
Qy 3096 GAGAAATGCTAATAACAAGTGGCCACATTTAAGAAAACAGGTGTATCAGGCCATGAGTT 3155
Db 2401 GAGAAATGCTAATAACAAGTGGCCACATTTAAGAAAACAGGTGTATCAGGCCATGAGTT 2460
Qy 3156 TATGAACCTAAGAGATGAATCACTGAAGACTTCAATATGTACTTTTATCATTTACTTCCAAA 3215
Db 2461 TATGAACCTAAGAGATGAATCACTGAAGACTTCAATATGTACTTTTATCATTTACTTCCAAA 2520
Qy 3216 ACCCAGCATAGCAGGCTGAAATATGTCAGAGAGAAAGAGAAAAACAAGAAACAAAGAT 3275
Db 2521 ACCCAGCATAGCAGGCTGAAATATGTCAGAGAGAAAGAGAAAAACAAGAAACAAAGAT 2580
Qy 3276 GAAGCAATTCGCGCCACACAGCTCCTGAAATTCGCTGCTGCTTTTTCAGCAAGTGAATTAAC 3335
Db 2581 GAAGCAATTCGCGCCACACAGCTCCTGAAATTCGCTGCTGCTTTTTCAGCAAGTGAATTAAC 2640
Qy 3336 CTTCTCAATGTGATATCATGATGTACATTTCTCAGACCGTATTTTGGCGGCAATAGAC 3395
Db 2641 CTTCTCAATGTGATATCATGATGTACATTTCTCAGACCGTATTTTGGCGGCAATAGAC 2700
Qy 3396 ACAGATTTCTAACTTGTGGACCGAAGGATGTCTCCAAATGGCTTTTTCATATCTGGCATTTG 3455
Db 2701 ACAGATTTCTAACTTGTGGACCGAAGGATGTCTCCAAATGGCTTTTTCATATCTGGCATTTG 2760

Qy 3456 GGTTTACTAGAGAGAAAGCAACAGCTTCAAAAAGCTCTCTGAAGAAAGTAAACATTTTGAC 3515
Db 2761 GGTTTACTAGAGAGAAAGCAACAGCTTCAAAAAGCTCTCTGAAGAAAGTAAACATTTTGAC 2820
Qy 3516 TTTTATCATAAAGCTTCAAGATTGGGAAGTTTCAGCCATGAATATACAAATGCTTTTGGAA 3575
Db 2821 TTTTATCATAAAGCTTCAAGATTGGGAAGTTTCAGCCATGAATATACAAATGCTTTTGGAA 2880
Qy 3576 AAACCTAAAGGAATTTCCCAAGTTAGAAGGCCAGAGGACATGATAACGTGATACTTTCAG 3635
Db 2881 AAACCTAAAGGAATTTCCCAAGTTAGAAGGCCAGAGGACATGATAACGTGATACTTTCAG 2940
Qy 3636 ATGTTTTCACACAGTGAAGCGAATTAAGAGAAAAATCTTGTGTTTAAATTTGTAGCAACACATCA 3695
Db 2941 ATGTTTTCACACAGTGAAGCGAATTAAGAGAAAAATCTTGTGTTTAAATTTGTAGCAACACATCA 3000
Qy 3696 GGATCGGAATCTTATTAAGAAATGATGAGATTAATCTCATGATAAAGAAAAAGCAGAAAC 3755
Db 3001 GGATCGGAATCTTATTAAGAAATGATGAGATTAATCTCATGATAAAGAAAAAGCAGAAAC 3060
Qy 3756 AGAAAAGCTGAAGCTGTAGGCTACATCGCCAGAAAGATCATGGCTCAGATGTCTGCCTTA 3815
Db 3061 AGAAAAGCTGAAGCTGTAGGCTACATCGCCAGAAAGATCATGGCTCAGATGTCTGCCTTA 3120
Qy 3816 CAGAAAATCTTCAATTTGAAAACCTCATAAACTCATGTATGCAATATCATCAGAAATGCTCTGG 3875
Db 3121 CAGAAAATCTTCAATTTGAAAACCTCATAAACTCATGTATGCAATATCATCAGAAATGCTCTGG 3180
Qy 3876 AAAGAGATTTCCATTTATGGAGGAAGAGAGCAACCCAGCAGTCACTAGTACTCTAGAAAT 3935
Db 3181 AAAGAGATTTCCATTTATGGAGGAAGAGAGCAACCCAGCAGTCACTAGTACTCTAGAAAT 3240
Qy 3936 GCTTTGGGTCTTAAACGGGTCTCATCTGTTACTGAAAAGAGAGGTCTGACGTGCATCTT 3995
Db 3241 GCTTTGGGTCTTAAACGGGTCTCATCTGTTACTGAAAAGAGAGGTCTGACGTGCATCTT 3300
Qy 3996 TGCCAGAGAAACAGGAGGTGAAAATAGAAATAATGCAATGATATCGGCTGTGTC 4055
Db 3301 TGCCAGAGAAACAGGAGGTGAAAATAGAAATAATGCAATGATATCGGCTGTGTC 3360
Qy 4056 CAGAAAATCTACTGCTTTAAACCCAGCAGAGGGGAAAAACCCATAGAACTCTCAGGAGAAAGCC 4115
Db 3361 CAGAAAATCTACTGCTTTAAACCCAGCAGAGGGGAAAAACCCATAGAACTCTCAGGAGAAAGCC 3420
Qy 4116 CTAGACCCACTTTTTCATGATGATCAGACTTGCGCATATGGAATTTATACAGGAAGCTGTGCT 4175
Db 3421 CTAGACCCACTTTTTCATGATGATCAGACTTGCGCATATGGAATTTATACAGGAAGCTGTGCT 3480
Qy 4176 CATGTAATGACAGCAGTGTGCTGCGCAGAAATATTTTGAAGCTGTACAGCTCAGCTCTCAG 4235
Db 3481 CATGTAATGACAGCAGTGTGCTGCGCAGAAATATTTTGAAGCTGTACAGCTCAGCTCTCAG 3540
Qy 4236 CAGCGCATTTCAATGTTGACCTTTTGGACTTGGAAAGTGGAGAAATATCTTTGGCCCTCTTTGC 4295
Db 3541 CAGCGCATTTCAATGTTGACCTTTTGGACTTGGAAAGTGGAGAAATATCTTTGGCCCTCTTTGC 3600
Qy 4296 AAATCTCTGTGCAATATCTGATCCCAATTTATTTCTTTGCACTCTCAAAAGATTAACAGT 4355
Db 3601 AAATCTCTGTGCAATATCTGATCCCAATTTATTTCTTTGCACTCTCAAAAGATTAACAGT 3660
Qy 4356 GAGAAATGCAAGTCTCTTGTCTCAACTTTTGAACCTTGGCAGCTGGATACAGACTGTTCTG 4415
Db 3661 GAGAAATGCAAGTCTCTTGTCTCAACTTTTGAACCTTGGCAGCTGGATACAGACTGTTCTG 3720
Qy 4416 GCCAGAAATATCAGGTTTAAATATTAAGACATGCTTAAAGGAGAAAAACCCAAATTTCTTATTTTC 4475
Db 3721 GCCAGAAATATCAGGTTTAAATATTAAGACATGCTTAAAGGAGAAAAACCCAAATTTCTTATTTTC 3780
Qy 4476 TTTTAAATCAGGAATGGGAGATTTCTACTTTGAGTTTCCATTTCCATCTCTGAGTTTGGCGTT 4535
Db 3781 TTTTAAATCAGGAATGGGAGATTTCTACTTTGAGTTTCCATTTCCATCTCTGAGTTTGGCGTT 3840

```
QY 4536 GAGTCTTCGATTAAATATTCABAATAGCATCAAGGAATGGTTATTCCTCTTTGCCACAACA 4595
Db 3841 GAGTCTTCGATTAAATATTCABAATAGCATCAAGGAATGGTTATTCCTCTTTGCCACAACA 3900
QY 4596 ATTTATAGAAATGGATTGAAGAGGCCACCTGATGAAGAGGATCCTCAGTCCCCATGCTG 4655
Db 3901 ATTTATAGAAATGGATTGAAGAGGCCACCTGATGAAGAGGATCCTCAGTCCCCATGCTG 3960
QY 4656 ACTTGAGACCTGCGCTTCTACTATCCAGGCAATTTGAAAATCTATTTGGAGATGAAGGA 4715
Db 3961 ACTTGAGACCTGCGCTTCTACTATCCAGGCAATTTGAAAATCTATTTGGAGATGAAGGA 4020
QY 4716 ARACTCTCTGTTGGAGCACTTCAAAATAGGAGCATAATGGTCTGAAGCAATTAATGCAG 4775
Db 4021 ARACTCTCTGTTGGAGCACTTCAAAATAGGAGCATAATGGTCTGAAGCAATTAATGCAG 4080
QY 4776 TTTGCAAGTTGCACAGAGGATTAACCTGCTCAGGTCCTGATACAGAAACATCTGGTTCCG 4835
Db 4081 TTTGCAAGTTGCACAGAGGATTAACCTGCTCAGGTCCTGATACAGAAACATCTGGTTCCG 4140
QY 4836 CTTCTATCAGTTGTTCTTCTTAACATAAATAAGAAATACAGATGACCATGCTTCTGTTCTATA 4895
Db 4141 CTTCTATCAGTTGTTCTTCTTAACATAAATAAGAAATACAGATGACCATGCTTCTGTTCTATA 4200
QY 4896 GATCTGTTTCATGTTTGGTGGTCTGTGTAGCATTCCTCATCTTGTATTGGATGAC 4955
Db 4201 GATCTGTTTCATGTTTGGTGGTCTGTGTAGCATTCCTCATCTTGTATTGGATGAC 4260
QY 4956 CTTGTTGATCTGAGGCTTCTTTCAGTTAGTTCTTCTCTATAACCACTTTTATCTTCCAT 5015
Db 4261 CTTGTTGATCTGAGGCTTCTTTCAGTTAGTTCTTCTCTATAACCACTTTTATCTTCCAT 4320
QY 5016 TTGATCACCATGGCACACATGCTTTCAGATTACTTACAGTAGACACAGGCTTACCCCTT 5075
Db 4321 TTGATCACCATGGCACACATGCTTTCAGATTACTTACAGTAGACACAGGCTTACCCCTT 4380
QY 5076 GCTCAGGTTCAAGAGACAGTAGAGGCTTCAATCCGATCTCTTTCTTTGAGAAATTT 5135
Db 4381 GCTCAGGTTCAAGAGACAGTAGAGGCTTCAATCCGATCTCTTTCTTTGAGAAATTT 4440
QY 5136 TCTCAATATACAAAGTGCTCCATTTGGTGTGATATTCCTGGCTGGTATTGTGGGTCTCA 5195
Db 4441 TCTCAATATACAAAGTGCTCCATTTGGTGTGATATTCCTGGCTGGTATTGTGGGTCTCA 4500
QY 5196 CTGAAGAATGGCATCACCCCTTATCTTCGCTGTGCTGCAATGTTTTTCCACTATTTACTT 5255
Db 4501 CTGAAGAATGGCATCACCCCTTATCTTCGCTGTGCTGCAATGTTTTTCCACTATTTACTT 4560
QY 5256 GGGGTAACTCCGCTGAGGAACCTGCATACCAATTTCTGCAAGAGGAGATACAGTGCACTC 5315
Db 4561 GGGGTAACTCCGCTGAGGAACCTGCATACCAATTTCTGCAAGAGGAGATACAGTGCACTC 4620
QY 5316 TGTAGTATCTATCTTACTCAAAATTTGTTCTGCTCTCCAGGAATTTGGGATCT 5375
Db 4621 TGTAGTATCTATCTTACTCAAAATTTGTTCTGCTCTCCAGGAATTTGGGATCT 4680
QY 5376 GTAAGGCCCTTGTCTCAGAGGTGGTGTGAGATCTCTGCTTACTAACTGTTTGAAGCAA 5435
Db 4681 GTAAGGCCCTTGTCTCAGAGGTGGTGTGAGATCTCTGCTTACTAACTGTTTGAAGCAA 4740
QY 5436 AAAAAACACCGTGTGAGTACCTTAGAAAAAGAAATAGTTTGAATAGAGCTTCTGATGAC 5495
Db 4741 AAAAAACACCGTGTGAGTACCTTAGAAAAAGAAATAGTTTGAATAGAGCTTCTGATGAC 4800
QY 5496 TATAGTGTCTCTGATCAAGCTTCTCAATTTGAGTGGCCACGCTCTGAGATGATGAG 5555
Db 4801 TATAGTGTCTCTGATCAAGCTTCTCAATTTGAGTGGCCACGCTCTGAGATGATGAG 4860
QY 5556 CGAAGACATCTGTCTCTGCTCTTTCTGTGGGCTTACTATGTTCTCAGAACATTTGC 5615
Db 4861 CGAAGACATCTGTCTCTGCTCTTTCTGTGGGCTTACTATGTTCTCAGAACATTTGC 4920
QY 5616 TCCAGGAAATTTGTGAAACGGGGAAGAGGTTGGAGCTTGCAATTTTTCACGCACTTCACTGT 5675
```

```
Db 4921 TCCAGGAAATTTGTGAACGGGAGAGGTTGGAGCTTGCAATTTTTCACCACTTCACTGT 4980
QY 5676 GGAGCCGGAGTCTGCATTTTCTCTAAAAATCAGAGAATGCCAGTGGTCTCTGGTTGAAGGT 5735
Db 4981 GGAGCCGGAGTCTGCATTTTCTCTAAAAATCAGAGAATGCCAGTGGTCTCTGGTTGAAGGT 5040
QY 5736 AAAGCCAGAGGCTGTGCTTATCCAGTCTCTTACTTGGATGAATATGAGAAAACAGACCTT 5795
Db 5041 AAAGCCAGAGGCTGTGCTTATCCAGTCTCTTACTTGGATGAATATGAGAAAACAGACCTT 5100
QY 5796 GGCCTGAAGAGGCGCAACCCCTTCTATTTATCTCTGTAGCGGTATCGGAAGCTTCCATTG 5855
Db 5101 GGCCTGAAGAGGCGCAACCCCTTCTATTTATCTCTGTAGCGGTATCGGAAGCTTCCATTG 5160
QY 5856 GTCTGGCAACCACTGCAATTAAGAAGATTTGCTAGGAGCCAAAGAGACTAATCAGATG 5915
Db 5161 GTCTGGCAACCACTGCAATTAAGAAGATTTGCTAGGAGCCAAAGAGACTAATCAGATG 5220
QY 5916 TTATTTGGATTCACTGGCAGTTACTGTGA 5945
Db 5221 TTATTTGGATTCACTGGCAGTTACTGTGA 5250

RESULT 3
AR487565 5205 bp DNA linear PAR 14-MAY-2004
DEFINITION Sequence 18 from patent US 6706505.
ACCESSION AR487565
VERSION AR487565.1 GI:47252815
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 5205)
AUTHORS Han, H.-Q. and Kwak, K.
TITLE Human E3.alpha. ubiquitin ligase family
JOURNAL Patent: US 6706505-A 18 16-MAR-2004;
FEATURES Location/Qualifiers
1..5205
source /organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 81.4%; Score 5136; DB 6; Length 5205;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 5189; Conservative 10; Mismatches 6; Indels 45; Gaps 1;

QY 696 ATGCGGACGAGGAGGCTGAGGTACTGAGAGGATGGAATCAGCGCGGAGTTACCCAG 755
Db 1 ATGCGGACGAGGAGGCTGAGGTACTGAGAGGATGGAATCAGCGCGGAGTTACCCAG 60
QY 756 ACCCTTCAGGCTCTGGCATCTTGGTGGATCAGCAAGTTGATTTTATCTGCTTCTTG 815
Db 61 ACCCTTCAGGCTCTGGCATCTTGGTGGATCAGCAAGTTGATTTTATCTGCTTCTTG 120
QY 816 CATCATTTGGCAAAATTTGGTGCAGAAATTTACTTTGCTGAAATGGACCAGACTTGGAA 875
Db 121 CATCATTTGGCAAAATTTGGTGCAGAAATTTACTTTGCTGAAATGGACCAGACTTGGAA 180
QY 876 AAGCAGGAGAAAGTGTAACAATGCAATTTACTCTCACTCGAATGGTACTTATTGGA 935
Db 181 AAGCAGGAGAAAGTGTAACAATGCAATTTACTCTCACTCGAATGGTACTTATTGGA 240
QY 936 GAAGATCCAGATTTTGTCTTAGAGAAATTTGAGACAGTGGAGCATTTTCAGCTTTGTGG 995
Db 241 GAAGATCCAGATTTTGTCTTAGAGAAATTTGAGACAGTGGAGCATTTTCAGCTTTGTGG 300
QY 996 AGGGTTTTCAAAAGTGGAGAGACCACTTATTTCTTCAGGGATTTGTGCAATTCACCA 1055
Db 301 AGGGTTTTCAAAAGTGGAGAGACCACTTATTTCTTCAGGGATTTGTGCAATTCACCA 360
QY 1056 TGTGTACTCTGTATGGAATGTCTTCCAGGACAGTGTTCATAAAAAATCATCGTTACAAGATG 1115
```


QY 3276 GAAGCAATTCGCCGCCACCACCTCTCTGAAATTCGCCCCCTGCTTTCAGGCAAGTGATTAAC 3335
DB 2536 GAAGCAATTCGCCGCCACCACCTCTCTGAAATTCGCCCCCTGCTTTCAGGCAAGTGATTAAC 2595
QY 3336 CTTCCTCAACTGTGATATCATGATGTACATCTCTCAGGACCGTATTTGAGCGGCAATAGAC 3395
DB 2596 CTTCCTCAACTGTGATATCATGATGTACATCTCTCAGGACCGTATTTGAGCGGCAATAGAC 2655
QY 3396 ACAGATTTCTAACTTGTGGACCGAAGGGATGCTCCAAATGGCTTTTCATATCTCGCATTTG 3455
DB 2656 ACAGATTTCTAACTTGTGGACCGAAGGGATGCTCCAAATGGCTTTTCATATCTCGCATTTG 2715
QY 3456 GGTTTACTAGAGAGAGCAACAGCTTCAAAGCTCTCGAAGAGTAAAGTAAAGTAAAGTAAAG 3515
DB 2716 GGTTTACTAGAGAGAGCAACAGCTTCAAAGCTCTCGAAGAGTAAAGTAAAGTAAAGTAAAG 2775
QY 3516 TTTTATCATAGGCTTCAAGATTTGGAGTTGGAGTTGAGGATGCTCAGCCATGATATCAAAATGCTTTTGGAA 3575
DB 2776 TTTTATCATAGGCTTCAAGATTTGGAGTTGAGGATGCTCAGCCATGATATCAAAATGCTTTTGGAA 2835
QY 3576 AAACCTCAAAGGAATTTCCCGAGTTAGAGGCCAGAGGACATGATTAACGTGGATATCTTCAG 3635
DB 2836 AAACCTCAAAGGAATTTCCCGAGTTAGAGGCCAGAGGACATGATTAACGTGGATATCTTCAG 2895
QY 3636 ATGTTTGAACAGTGAAGCGATTAAAGAGAAATCTTGTTTAAATTTGTAGCAACCAATCA 3695
DB 2896 ATGTTTGAACAGTGAAGCGATTAAAGAGAAATCTTGTTTAAATTTGTAGCAACCAATCA 2955
QY 3696 GGATCGGAATCTTAAAGATGATCAGATTAATCTCATGATTAAGAAAGAGCAAGCAAAA 3755
DB 2956 GGATCGGAATCTTAAAGATGATCAGATTAATCTCATGATTAAGAAAGAGCAAGCAAAA 3015
QY 3756 AGAAAGCTGAAGCTGTAGGCTACATCGCCAGAGAGATCATGGCTCAGATGTCTGCCCTTA 3815
DB 3016 AGAAAGCTGAAGCTGTAGGCTACATCGCCAGAGAGATCATGGCTCAGATGTCTGCCCTTA 3075
QY 3816 CAGAAAACTTCAATTTGAACTCATAACTCATGATATGACATACATCAGAAATGCTCGG 3875
DB 3076 CAGAAAACTTCAATTTGAACTCATAACTCATGATATGACATACATCAGAAATGCTCGG 3135
QY 3876 AAGGAAGATTCATTTATGGAG 3935
DB 3136 AAGGAAGATTCATTTATGGAG 3195
QY 3936 GCTTTGGGTCCTAAACGGGTCCTATCTGTACTGAAAGAGAGAGAGAGAGAGAGAGAGAG 3995
DB 3196 GCTTTGGGTCCTAAACGGGTCCTATCTGTACTGAAAGAGAGAGAGAGAGAGAGAGAGAG 3255
QY 3996 TGCCAAAG 4055
DB 3256 TGCCAAAG 3315
QY 4056 CAGAAATCTAGTCCCTTAACCCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4115
DB 3316 CAGAAATCTAGTCCCTTAACCCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3375
QY 4116 CTAGACCCACTTTTTCATGGATCCAGACTTGGCATATGAACTTATACAGGAGAGAGAGAG 4175
DB 3376 CTAGACCCACTTTTTCATGGATCCAGACTTGGCATATGAACTTATACAGGAGAGAGAGAG 3435
QY 4176 CATGTAATGACGAG 4235
DB 3436 CATGTAATGACGAG 3495
QY 4236 CAGCGCATTCATGTTGACCTTTTGAATTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4295
DB 3496 CAGCGCATTCATGTTGACCTTTTGAATTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3555
QY 4296 AAATCTGTGCAATCTGATCCCATTTATCTTTGCAACCTTCAAAAGATTAACAGT 4355
DB 3556 AAATCTGTGCAATCTGATCCCATTTATCTTTTGCACCTTCAAAAGATTAACAGT 3615

QY 4356 GAGATGACGATGCTCTTGTCTCAACTTTTGGACCTTGGCAGCGTGATACAGACTGTTCTG 4415
DB 3616 GAGATGACGATGCTCTTGTCTCAACTTTTGGACCTTGGCAGCGTGATACAGACTGTTCTG 3675
QY 4416 GCCAGATATCAGGTTTAAATATATAGACATGCTAAAGGAGAGAAACCCAAATCTCTATTTTC 4475
DB 3676 GCCAGATATCAGGTTTAAATATATAGACATGCTAAAGGAGAGAAACCCAAATCTCTATTTTC 3735
QY 4476 TTTAATCAAGGAATGGGAGATTTCTACTTTGGAGTTCCATTCATCTCCTGAGTTTGGCGTT 4535
DB 3736 TTTAATCAAGGAATGGGAGATTTCTACTTTGGAGTTCCATTCATCTCCTGAGTTTGGCGTT 3795
QY 4536 GAGCTTCGATTAATATTTCAATAGCATCAAGGAATGGTTATTTCTCTTGGCCACAACA 4595
DB 3796 GAGCTTCGATTAATATTTCAATAGCATCAAGGAATGGTTATTTCTCTTGGCCACAACA 3855
QY 4596 ATTTATAGAAATGGATTTGAAAGTGCCTCATGATGAAGGGATCTCTCGAGTCCCATGCTG 4655
DB 3856 ATTTATAGAAATGGATTTGAAAGTGCCTCATGATGAAGGGATCTCTCGAGTCCCATGCTG 3915
QY 4656 ACCTGGAGCACCCTGCGCTTTCACTATCCAGGCAATTTGAAATCTATTGGGAGATGAAGA 4715
DB 3916 ACCTGGAGCACCCTGCGCTTTCACTATCCAGGCAATTTGAAATCTATTGGGAGATGAAGA 3975
QY 4716 AAACTCTGTTTGGAGCATCTTCAAAAATAGGAGAGATTAATGGTCTGAAAGCATTAATGAG 4775
DB 3976 AAACTCTGTTTGGAGCATCTTCAAAAATAGGAGAGATTAATGGTCTGAAAGCATTAATGAG 4035
QY 4776 TTTGAGTTGACAGAGAGATTTACCTGCTCAGGTCCTGATACAGAAACATCTGCTTCTG 4835
DB 4036 TTTGAGTTGACAGAGAGATTTACCTGCTCAGGTCCTGATACAGAAACATCTGCTTCTG 4095
QY 4836 CTCTCTCAGTTGTTCTTCTTAACATATAAATCAGAGATACACCATGCTTCTGCTCTATA 4895
DB 4096 CTCTCTCAGTTGTTCTTCTTAACATATAAATCAGAGATACACCATGCTTCTGCTCTATA 4155
QY 4896 GATCTGTTTCAATGTTTGGTGGTCTGTTAGCATTTCCCATCTCTGTTATTTGGAGTAC 4955
DB 4156 GATCTGTTTCAATGTTTGGTGGTCTGTTAGCATTTCCCATCTCTGTTATTTGGAGTAC 4215
QY 4956 CCTGTTGATCTGACGCTTCTTCTGAGTTAGTTCTCTTATACCACTTATCTCTTCCAT 5015
DB 4216 CCTGTTGATCTGACGCTTCTTCTGAGTTAGTTCTCTTATACCACTTATCTCTTCCAT 4275
QY 5016 TTGATCACTACCTGACACATGCTTTCAGATATCTACTTACAGTAGACACAGGCTTACCCCTT 5075
DB 4276 TTGATCACTACCTGACACATGCTTTCAGATATCTACTTACAGTAGACACAGGCTTACCCCTT 4335
QY 5076 GCTCAGGTTCAAG 5135
DB 4336 GCTCAGGTTCAAG 4395
QY 5136 TCTCAATATACAGTGGTCCATTTGGGTTGATATCTCTGCTGCTGCTTATTTGAGGCTTCA 5195
DB 4396 TCTCAATATACAGTGGTCCATTTGGGTTGATATCTCTGCTGCTGCTTATTTGAGGCTTCA 4455
QY 5196 CTGAAGAGATGCACTACCCCTTATCTTCTGCTGCTGCTGCTTATTTTCCACTATTTACTT 5255
DB 4456 CTGAAGAGATGCACTACCCCTTATCTTCTGCTGCTGCTGCTTATTTTCCACTATTTACTT 4515
QY 5256 GGGGTAACTCCGCTGAGGAACTGCATACCAATTTCTGAGAGAGAGAGAGAGAGAGAGAG 5315
DB 4516 GGGGTAACTCCGCTGAGGAACTGCATACCAATTTCTGAGAGAGAGAGAGAGAGAGAGAG 4575
QY 5316 TGTAGCTATCTATCTTCTTACCTACAAATTTGTTTCTGCTCTTCCAGGATATTTGGAGTACT 5375
DB 4576 TGTAGCTATCTATCTTCTTACCTACAAATTTGTTTCTGCTCTTCCAGGATATTTGGAGTACT 4635
QY 5376 GTAAAGGCTTGTCTCCAG 5435
DB 4636 GTAAAGGCTTGTCTCCAG 4695
QY 5436 AAAAACAACCGTGTGAGTACCCCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5495

[illegible]

RESULT 4
 AF525401 LOCUS
 DEFINITION Homo sapiens UBRI E3a ligase mRNA, partial cds.
 ACCESSION AF525401
 VERSION AF525401.1 GI:27451603
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 5128)
 REFERENCE
 AUTHORS Dgany,O., Avidan,N., Delaunay,J., Krasnov,T., Shalmon,L.,
 Shalev,H., Eidelits-Markus,T., Kapelushnik,J., Cattán,D.,
 Pariente,A., Tulliez,M., Cretien,A., Schischmanoff,P.O.,
 Iolascon,A., Fibach,E., Koren,A., Rossler,J., Le Merer,M.,
 Yaniv,I., Zaizov,R., Ben-Asher,E., Olender,T., Lancet,D.,
 Beckmann,J.S. and Tamary,H.
 CONGENITAL DYSERYTHROPOIETIC ANEMIA TYPE I IS CAUSED BY MUTATIONS
 IN CODANIN-1
 Am. J. Hum. Genet. 71 (6), 1467-1474 (2002)
 22340442
 MEDLINE 12434312
 PUBMED
 2 (bases 1 to 5128)
 REFERENCE
 AUTHORS Ben-Asher,E.
 TITLE Direct Submission
 JOURNAL Submitted (27-JUN-2002) Molecular Genetics, The Weizmann Institute
 of Science and the Schneider Medical Center, Hertzl Street,
 Rehovot 76100, Israel
 FEATURES
 source
 1. 5128
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="15"
 /map="15q13"
 /cell_type="erythroid"

```

CDS
<1..5128
/codon_start=1
/product="UBR1 E3a ligase"
/protein_id="AAO14997.1"
/db_xref="db:27451604"
/tranlation="*LPQTPORLASWDDQVDFYTAFLHLAQLVPEIYFAEMDPDLERK
QESVMSIPTPLEWFLNFGEDPDI CLEKLKHSAGFOLCGRFPSKSGTYSICRDCAIDP
TCVLCMDCFQDSVHKHNRKMYKHTSTGGGFCDCGDTAWKTPFCFVNHEPRGRTAIKEN
SCPLNEEVIQARKIPFVQIVYVEMTIIWEERKELPPELOIRKRNRYVCVLFNDEH
HSYDHYISQALDCLEAQLHTTAIDKEGRRAVKAGAYACQAKEDIKSHSNV
SOHLPEVHLHSEIMAHQKILRLGSMWNKIMSYSDDFQIICQACALREBPENFCL
ISRLMDWAKLYKAGARKILHLLIFSSPFMEMEYKKLIFAMEFVKYKQLOKEYISDDDH
RSISITALSVMQPTPTLARHLLIEQNVISVITELLEVLPYELDRNNKNFQGYSDQ
KUGRVAIVADUKYILISIKPTIWIERLRFQFLEGRFSFLKLTFCMQMEIREKQVGH
I EYDPDWEAIIAQMQLNKLMFQEWCADEELLVAYKGECHKAVCMSTSPISSK
TVVYSGHSLSTKYSRVSDVLNHLPSRLTLAGLHVRLSRGLGASRLHEFVGFDFQ
FVVLVYPCRLCLVAQVAVEMRRRLSLISQVFFYQDVKRCSEVMDKDIIMLOIGA
SLMDPNKFLIIVLORVELAAFNKTIISTKDQDI IKOYNTLVBEMLOVLIIVGERVVP
GVGNVTEKEVWTRELIHLLICIBPMHSAIAKNLPENENNETGLENVKATPKKGV
SGHGVYELNDEKIDFNTFYHKTQHSKAEMHQKRRQENKQDASLPFPPPEFCF
APSKVNTLQDILMYLIRTVFERAIDTDSNLWTEGMLQWAFHILAGLLEEKQLOK
APBEETVTFYTKASRLGSSAMNIMLLEKIGIPQLEGQKTMITWLOMFDTVFKRLR
EKSLIVATTSGESIKNDIEIYHDKBAKRKAZEARLHROKIMQAMSLAQNLCTET
HKLMYNTSEMPKEDSIWBEESTPAVDSYRIALGPGRGPSVTEKEVLCTILCOREQ
EYKIENNAVLSVCQKSTALQHRGKPIELSGEALDPLMPDPLGAYGTGSGCHVM
HAVQYIFEAQVLSQQRIHVLFDLESGLYCLPCKSLCNTVFIILPQPKINSE
NADALAQLITLARIWITGLIARISGVNIHRAKGENPIIPFQMGMSDLEFHLISLSPG
VRESSIKYSIKEMVILFPAITYRIGLVPPDPERVPMLTWSTCAFTQATENLGG
DEKGPLGALQNRHGLKALMPQVAQRIITCPQVLIQKHLVRLLSVLPNIKSEDTPT
CULSIDLPHVLQVDSHPSLVWDDPVDLPQPSVSSYNHLVFLHITWAMLMQIILLT
VTGPLPLAQVEDSEAHASASFPARIQSYTSGSIGCDIPGMVYLSKNGLITPVLRC
AALFFHYLGVTPPEELHTNSAGEYSALCSYLSLPXNLPFLPQEWDTVRPLLRWC
ADPALLNCQKNTVVRYPKRNLSIELPDYSLNQAQSHFPCRSADDERHVPILC
LPGATILCSQNTCCQBIIVNGEYVAGACIFHALHCGAGVCIFLKIRECRVVLVEGKARG
AYPAPYLDEYGETDPLKRGKNPLHFLVSGTEAPLVRQHQCI"
ORIGIN

Query Match      80.6%; Score 5085.6; DB 9; Length 5128;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 5121; Conservative 1; Mismatches 6; Indels 3; Gaps 3;

Qy 747 TTATCCCGAGACCCTCAGCGCTCTGGCATCTTGTTGGGATCAGCAAGTTGATTTTATACT 806
Db 1 TTATCCCGAGACCCTCAGCGCTCTGGCATCTTGTTGGGATCAGCAAGTTGATTTTATACT 60

Qy 807 GCTTTCTTGCAATCATTTGGCAATTTGGTGCCAGAAATTTATCTTGCTGAAATGGACCCA 866
Db 61 GCTTTCTTGCAATCATTTGGCAATTTGGTGCCAGAAATTTATCTTGCTGAAATGGACCCA 120

Qy 867 GACTTGGAAAAGCAGGAGGAAGTGTACAAATGTCATAATTTCACTCCACTGGAAATGGTAC 926
Db 121 GACTTGGAAAAGCAGGAGGAAGTGTACAAATGTCATAATTTCACTCCACTGGAAATGGTAC 180

Qy 927 TTATTTTGGAGAAGATCCAGATATTTGCTTAGAGAAATTTGAAGCACAGTGCGAGCATTTTCAG 986
Db 181 TTATTTTGGAGAAGATCCAGATATTTGCTTAGAGAAATTTGAAGCACAGTGCGAGCATTTTCAG 240

Qy 987 CTTTGTGGGAGGGTTTTCAAAGTGGAGAGACAACCTATTTCTTGGAGGGGATTTGTGCAATT 1046
Db 241 CTTTGTGGGAGGGTTTTCAAAGTGGAGAGACAACCTATTTCTTGGAGGGGATTTGTGCAATT 300

Qy 1047 GATCCAAATGTTGACTCTGTATGGACTGTTCCAGGACAGTCTTTCATAAAAATCATCGT 1106
Db 301 GATCCAAATGTTGACTCTGTATGGACTGTTCCAGGACAGTCTTTCATAAAAATCATCGT 360

Qy 1107 TACAAGATGCATACTTCTACTCGGAGGAGGGTCTGTGCACTGTGGAGACACAGAGGCATGG 1166
Db 361 TACAAGATGCATACTTCTACTCGGAGGAGGGTCTGTGCACTGTGGAGACACAGAGGCATGG 420

Qy 1167 AAAACTGGCCCTTTTGTGTAAATCATGAACCTGGAGAGCAGGTACTATAAAGAGAAAT 1226
Db 421 AAAACTGGCCCTTTTGTGTAAATCATGAACCTGGAGAGCAGGTACTATAAAGAGAAAT 480

```


QY 1227 TCACGCTGTCGTTGAAATGAAGAGCTTAATGTGTCAGCCGAGAAATATTTCTTCAGTG 1286
DB 481 TCACGCTGTCGTTGAAATGAAGAGCTTAATGTGTCAGCCGAGAAATATTTCTTCAGTG 540
QY 1287 ATAAAATATGTCGTAGAATGACTATATGGGAAGAGGAAAAAGAACTGCTCTCAACTC 1346
DB 541 ATAAATATGTCGTAGAATGACTATATGGGAAGAGGAAAAAGAACTGCTCTCAACTC 600
QY 1347 CAGATAAGGGAGAAAAATGAAGATACTATATGTCCTTTTCAATGATGAACACCAATCA 1406
DB 601 CAGATAAGGGAGAAAAATGAAGATACTATATGTCCTTTTCAATGATGAACACCAATCA 660
QY 1407 TATGACCAAGCTATATACAGCCTACAAAGAGCTCTTGACTGTGAGCTGCGAGAGCCGAG 1466
DB 661 TATGACCAAGCTATATACAGCCTACAAAGAGCTCTTGACTGTGAGCTGCGAGAGCCGAG 720
QY 1467 TTGCATACCACTGCCATTGACAAAGAGGGTCTGCGGCTGTTAAAGCGGAGCTTATGCT 1526
DB 721 TTGCATACCACTGCCATTGACAAAGAGGGTCTGCGGCTGTTAAAGCGGAGCTTATGCT 780
QY 1527 GCTTCCAGGAAGCAAGGAAGATATAAGAGTCAATTCAGAAAAATGTCCTCAACATCCA 1586
DB 781 GCTTCCAGGAAGCAAGGAAGATATAAGAGTCAATTCAGAAAAATGTCCTCAACATCCA 840
QY 1587 CTTCACTAGAGATATACACTCAGAGATATGCTCATCAGAAATTTGCTTTGCGTCTT 1646
DB 841 CTTCACTAGAGATATACACTCAGAGATATGCTCATCAGAAATTTGCTTTGCGTCTT 900
QY 1647 GCTTCCAGGAAGCAAGGAAGATATAAGAGTCAATTCAGAAAAATGTCCTCAACATCCA 1706
DB 901 GCTTCCAGGAAGCAAGGAAGATATAAGAGTCAATTCAGAAAAATGTCCTCAACATCCA 960
QY 1707 GCATGCTTAGAGAGAACTGACTCGAGAAATCCCTGCTCTCATAGCAGGTTAAATGCTT 1766
DB 961 GCATGCTTAGAGAGAACTGACTCGAGAAATCCCTGCTCTCATAGCAGGTTAAATGCTT 1020
QY 1767 TGGGATGCAAGCTTTATAAGGTCGCGTAAGATCCCTTCATGAATGATCTTCAGCAGT 1826
DB 1021 TGGGATGCAAGCTTTATAAGGTCGCGTAAGATCCCTTCATGAATGATCTTCAGCAGT 1080
QY 1827 TTTTATTTAGGAGATGGAATACAAAACTCTTTGCTATGGAATTTGTAAGTATATATAA 1886
DB 1081 TTTTATTTAGGAGATGGAATACAAAACTCTTTGCTATGGAATTTGTAAGTATATATAA 1140
QY 1887 CAATGCGAGAAAGATATATCAGTGATGATCATGACAGAGATATCTCTATAACTGCACTT 1946
DB 1141 CAATGCGAGAAAGATATATCAGTGATGATCATGACAGAGATATCTCTATAACTGCACTT 1200
QY 1947 TCAGTTCAGATGTTTACTGTTCTTACTCTGGCTCGACATCTTTATTTGAAGAGCAGAAATGTT 2006
DB 1201 TCAGTTCAGATGTTTACTGTTCTTACTCTGGCTCGACATCTTTATTTGAAGAGCAGAAATGTT 1260
QY 2007 ATCTCTGTCATTTACTGAACTCTGCTAGAGTTTACTGAGTACTTGGACAGGAACTAAT 2066
DB 1261 ATCTCTGTCATTTACTGAACTCTGCTAGAGTTTACTGAGTACTTGGACAGGAACTAAT 1320
QY 2067 AAATTTCACTTCCAGGTTATAGCAGGACCAATTTGGGAAGAGTATATGTCAGTATATGT 2126
DB 1321 AAATTTCACTTCCAGGTTATAGCAGGACCAATTTGGGAAGAGTATATGTCAGTATATGT 1380
QY 2127 GACCTAAAGTATATCTGATCAGCAAAACCAATATGACAGAAAGATTAAGAAATGAG 2186
DB 1381 GACCTAAAGTATATCTGATCAGCAAAACCAATATGACAGAAAGATTAAGAAATGAG 1440
QY 2187 TTCCCTTGAAGTTTTCGATCTTTTGAAGATTTCTTACTGTCGAGGATGGAAGAA 2246
DB 1441 TTCCCTTGAAGTTTTCGATCTTTTGAAGATTTCTTACTGTCGAGGATGGAAGAA 1500
QY 2247 ATCCGAAGCAGGTTGGGCAACACATTTGAAGTGAATCTGATTTGGAGGCTGCCATGCT 2306
DB 1501 ATCCGAAGCAGGTTGGGCAACACATTTGAAGTGAATCTGATTTGGAGGCTGCCATGCT 1560
QY 2307 ATACAGATGCAATTTGAAGAAATATTTTACTCATGTTTCCAAGAGTGGTGCTTGTGATGAA 2366

DB 1561 ATACAGATGCAATTTGAAGAAATATTTTACTCATGTTTCCAAGAGTGGTGCTTGTGATGAA 1620
QY 2367 GAACTCTTACTTGTGGCTTTATAAAGATGTGTCAAAAGCTGTGATGAGGTGCGAGTACCAGT 2426
DB 1621 GAACTCTTACTTGTGGCTTTATAAAGATGTGTCAAAAGCTGTGATGAGGTGCGAGTACCAGT 1680
QY 2427 TTTATATCTAGTAGCAGACAGTAGTAGTACAATGCTGTGTGACATAGTGTGGAACAAAGTCC 2486
DB 1681 TTTATATCTAGTAGCAGACAGTAGTAGTACAATGCTGTGTGACATAGTGTGGAACAAAGTCC 1740
QY 2487 TACAGAGTATCTGAGGATCTTGTAAAGCATACATCTGCCACTCTCTAGGACCCCTTGTGT 2546
DB 1741 TACAGAGTATCTGAGGATCTTGTAAAGCATACATCTGCCACTCTCTAGGACCCCTTGTGT 1800
QY 2547 CTTCAATGTACGTTTAAAGCAGGCTGGTGCTGTCTTCAAGACTGCATGAATTTGTGTCTTTT 2606
DB 1801 CTTCAATGTACGTTTAAAGCAGGCTGGTGCTGTCTTCAAGACTGCATGAATTTGTGTCTTTT 1860
QY 2607 GAGGACTTTCAAGTAGAGGTACTAGTGAATATCTTTTAAAGACTGCATGAATTTGTGTCTTTT 2666
DB 1861 GAGGACTTTCAAGTAGAGGTACTAGTGAATATCTTTTAAAGACTGCATGAATTTGTGTCTTTT 1920
QY 2667 CAGGTTGTGCTGAGATGTCGCGAGAAATGGAATGCTCTTTTAAAGACTGCATGAATTTGTGTCTTTT 2726
DB 1921 CAGGTTGTGCTGAGATGTCGCGAGAAATGGAATGCTCTTTTAAAGACTGCATGAATTTGTGTCTTTT 1980
QY 2727 TACCAAGATGTTAAAGTGCAGAGAAATGATGATGAAGATATCATCATGCTTCAGATTT 2786
DB 1981 TACCAAGATGTTAAAGTGCAGAGAAATGATGATGAAGATATCATCATGCTTCAGATTT 2040
QY 2787 GGTGCATCTTTAAATGGATCCCAATAGATTTCTTGTACTGGTACTTTCAGAGGTATGAACTT 2846
DB 2041 GGTGCATCTTTAAATGGATCCCAATAGATTTCTTGTACTGGTACTTTCAGAGGTATGAACTT 2100
QY 2847 GCGGAGGCTTTTAAACAGACCATATCTACAAAGACCAAGATTTGATTTAAACATATAT 2906
DB 2101 GCGGAGGCTTTTAAACAGACCATATCTACAAAGACCAAGATTTGATTTAAACATATAT 2160
QY 2907 ACCTTAATAGAGAAATGCTTCAGGCTCTCATCTATATTTGGGTGAGCGTTATGTACT 2966
DB 2161 ACCTTAATAGAGAAATGCTTCAGGCTCTCATCTATATTTGGGTGAGCGTTATGTACT 2220
QY 2967 GGAGTGGAAATGTGACCAAGAGAGTCAATAGAGAGAAATCATTCACCTGCTTTGC 3026
DB 2221 GGAGTGGAAATGTGACCAAGAGAGTCAATAGAGAGAAATCATTCACCTGCTTTGC 2280
QY 3027 ATTGAACCCATGCCACACAGTGCCTTGCCTTAAATTTTACCTGAGAAATGAATGA 3086
DB 2281 ATTGAACCCATGCCACACAGTGCCTTGCCTTAAATTTTACCTGAGAAATGAATGA 2340
QY 3087 ACTGGCTTAGAGAAATGTCATAAACAAGTGGCCACATTTTAAAGAACCAAGGTGATCAGGC 3146
DB 2341 ACTGGCTTAGAGAAATGTCATAAACAAGTGGCCACATTTTAAAGAACCAAGGTGATCAGGC 2400
QY 3147 CATGGAGTTTATGAATCAATGAAGTGAATCACTGAAAGACTTCAATATGCTACTTTATCAT 3206
DB 2401 CATGGAGTTTATGAATCAATGAAGTGAATCACTGAAAGACTTCAATATGCTACTTTATCAT 2460
QY 3207 TACTCCAAAACCCAGCATAGCAGGCTGAAATATGCGAAGAAAGAGAGAAACAGAA 3266
DB 2461 TACTCCAAAACCCAGCATAGCAGGCTGAAATATGCGAAGAAAGAGAGAAACAGAA 2520
QY 3267 AACAAAGATGAAGCAATTTGCGGCCACCACTCTCTGAATTTCTGCTTCTTTCAGCAAA 3326
DB 2521 AACAAAGATGAAGCAATTTGCGGCCACCACTCTCTGAATTTCTGCTTCTTTCAGCAAA 2580
QY 3327 GTGATTAACCTTCTCACTGTGATATCATGATGATCACTTCTCAGGACCGTATTTGAGCGG 3386
DB 2581 GTGATTAACCTTCTCACTGTGATATCATGATGATCACTTCTCAGGACCGTATTTGAGCGG 2640
QY 3387 GCAATAGACACAGATTTCTAACTGTGAGCCGAGGATGCTCCAAATGGCTTTTTCATATTT 3446

Db 2641 GCATATAGACACAGATTCTTAACCTTGTGACCGAAGGATGCTCCAAATGGCTTTTTCATATT 2700
Qy 3447 CTGGCAATTTGGGTTTACTAGAGAGAGCAACAGCTTCAAAAGCTCCCTGAGAGAACTA 3506
Db 2701 CTGGCAATTTGGGTTTACTAGAGAGAGCAACAGCTTCAAAAGCTCCCTGAGAGAACTA 2760
Qy 3507 ACATTTGACTTTTATCATAAAGCTTCAAGATTGGGAAGTTTCAGGCCATGAATATACAAATG 3566
Db 2761 ACATTTGACTTTTATCATAAAGCTTCAAGATTGGGAAGTTTCAGGCCATGAATATACAAATG 2820
Qy 3567 CTTTTGAAAAAATCTCAAAGGAATCCCCAGTTTGAAGGCCAGAGGAATGATATAACGTGG 3626
Db 2821 CTTTTGAAAAAATCTCAAAGGAATCCCCAGTTTGAAGGCCAGAGGAATGATATAACGTGG 2880
Qy 3627 ATACTTCAGATGTTTTGACACAGTGAAGCGATTAAAGAGAAATCTTGTTTAAATTTGTAGCA 3686
Db 2881 ATACTTCAGATGTTTTGACACAGTGAAGCGATTAAAGAGAAATCTTGTTTAAATTTGTAGCA 2940
Qy 3687 ACCACATCAGGATCGGAATCTATTAAAGATGATGAGATTACTCATGATATAAAGAAAAAGCA 3746
Db 2941 ACCACATCAGGATCGGAATCTATTAAAGATGATGAGATTACTCATGATATAAAGAAAAAGCA 3000
Qy 3747 GAAAGAAAAAGAAAGCTGAGCTCTAGGCTACATCGCCAGAGAGATCATGGCTCAGATG 3806
Db 3001 GAAAGAAAAAGAAAGCTGAGCTCTAGGCTACATCGCCAGAGAGATCATGGCTCAGATG 3060
Qy 3807 TCTGCTTTACAGAAAAAATCTTCAATGAAACTCAATAACTCATGATATGACAAATACATCAGAA 3866
Db 3061 TCTGCTTTACAGAAAAAATCTTCAATGAAACTCAATAACTCATGATATGACAAATACATCAGAA 3120
Qy 3867 ATGCGCTGGGAAAGAGATTCATTAATGAGGAAGAGAGACACCCAGCAGTCACTGACTAC 3926
Db 3121 ATGCGCTGGGAAAGAGATTCATTAATGAGGAAGAGAGACACCCAGCAGTCACTGACTAC 3180
Qy 3927 TCTAGAAATTTGGTCTTGGGCTCTAAACGGGGTCCATCTGTTACTGAAAAAGAGGTGCTGACG 3986
Db 3181 TCTAGAAATTTGGTCTTGGGCTCTAAACGGGGTCCATCTGTTACTGAAAAAGAGGTGCTGACG 3240
Qy 3987 TGATCTTTGCGCAAGAGAACAGAGGTGAATAATGAGAAATTAATGCCATGGTATTATCG 4046
Db 3241 TGATCTTTGCGCAAGAGAACAGAGGTGAATAATGAGAAATTAATGCCATGGTATTATCG 3300
Qy 4047 GCCTGTGTCCAGAAATCTACTGCCCTTAACCCAGCAGCGGAAAAACCCATAGAACTCTCA 4106
Db 3301 GCCTGTGTCCAGAAATCTACTGCCCTTAACCCAGCAGCGGAAAAACCCATAGAACTCTCA 3360
Qy 4107 GGAGAGCCCTAGACCCCACTTTTCATGGATCCAGACTTGGCATATGGAACCTTATACAGAA 4166
Db 3361 GGAGAGCCCTAGACCCCACTTTTCATGGATCCAGACTTGGCATATGGAACCTTATACAGAA 3420
Qy 4167 AGCTGTGTCTAATGACAGCAGTGTGCTGGCAGAGATATTTTGAAGCTGTACAGCTG 4226
Db 3421 AGCTGTGTGTATGTAATGACAGCAGTGTGCTGGCAGAGATATTTTGAAGCTGTACAGCTG 3480
Qy 4227 AGCTCTCAGCAGCGCATTCATGTTGACCTTTTGTGCTTGAAGGTGGAGAAATCTTTTGC 4286
Db 3481 AGCTCTCAGCAGCGCATTCATGTTGACCTTTTGTGCTTGAAGGTGGAGAAATCTTTTGC 3540
Qy 4287 CCTCTTTGCAAACTCTGTGCAATATCTGTGATCCCACTTATTTCTTTTGCACCTCAAAAG 4346
Db 3541 CCTCTTTGCAAACTCTGTGCAATATCTGTGATCCCACTTATTTCTTTTGCACCTCAAAAG 3600
Qy 4347 ATAAACAGTGAGAAATGCGATGCTCTTGCTGAACTTTTGAACCTGGCAGCGTGGATACAG 4406
Db 3601 ATAAACAGTGAGAAATGCGATGCTCTTGCTGAACTTTTGAACCTGGCAGCGTGGATACAG 3660
Qy 4407 ACTGTTCTGGCCAGAAATATCAGGTTTATATAAGACATGCTTAAAGGAGAAACCCCAATT 4466
Db 3661 ACTGTTCTGGCCAGAAATATCAGGTTTATATAAGACATGCTTAAAGGAGAAACCCCAATT 3720
Qy 4467 CCTATTTTCTTTAAATCAAGGAATGGAGATTCCTACTTTGGAGTTCCATTTCCATCTCGAGT 4526
Db 3721 CCTATTTTCTTTAAATCAAGGAATGGAGATTCCTACTTTGGAGTTCCATTTCCATCTCGAGT 3780

Qy 4527 TTTGGCTTGAGTCTTCGATTAATAATTTCAATAGCATCAAGGAAATGGTTATTCTCTTT 4586
Db 3781 TTTGGCTTGAGTCTTCGATTAATAATTTCAATAGCATCAAGGAAATGGTTATTCTCTTT 3840
Qy 4587 GCCACAACAATTTATAGAAATTTGGATTGGAATGCACTGATGAAAGGATCTTCGAGTC 4646
Db 3841 GCCACAACAATTTATAGAAATTTGGATTGGAATGCACTGATGAAAGGATCTTCGAGTC 3900
Qy 4647 CCATGCTGACCTGGAGCAGCTGCGCTTCACTATCCAGGCAATTTGAAAATCTATTGGGA 4706
Db 3901 CCATGCTGACCTGGAGCAGCTGCGCTTCACTATCCAGGCAATTTGAAAATCTATTGGGA 3960
Qy 4707 GATGAAGGAAACCTCTGTTTGGAGCACTTCAAAATAGGAGCATTAATGGTCTCAAAGCA 4766
Db 3961 GATGAAGGAAACCTCTGTTTGGAGCACTTCAAAATAGGAGCATTAATGGTCTCAAAGCA 4020
Qy 4767 TTAATGCAAGTTTGCAGTTGCAAGAGGATTAACCTGCTCTCAGGCTCTGATACAGAAACAT 4826
Db 4021 TTAATGCAAGTTTGCAGTTGCAAGAGGATTAACCTGCTCTCAGGCTCTGATACAGAAACAT 4080
Qy 4827 CTGGTTGCTCTTCTATCAGTTGTTCTTCTTAACATAAATCAGAAAGATACACCATGCCCTT 4886
Db 4081 CTGGTTGCTCTTCTATCAGTTGTTCTTCTTAACATAAATCAGAAAGATACACCATGCCCTT 4140
Qy 4887 CTGCTATAGATCTGTTTCAATGTTTGGTGGTCTGTTAGCAATTCCTATCCTTTGAT 4946
Db 4141 CTGCTATAGATCTGTTTCAATGTTTGGTGGTCTGTTAGCAATTCCTATCCTTTGAT 4200
Qy 4947 TGGGATGACCTGTTGATCTGACGCTTCTTCACTAGTTCTTCTCTATAAACACCTTTAT 5006
Db 4201 TGGGATGACCTGTTGATCTGACGCTTCTTCACTAGTTCTTCTCTATAAACACCTTTAT 4260
Qy 5007 CTCTTCCATTTGATCACCATGGGCACATGCTTCACTACTACTTACAGTATAGACACAGGC 5066
Db 4261 CTCTTCCATTTGATCACCATGGGCACATGCTTCACTACTACTTACAGTATAGACACAGGC 4320
Qy 5067 CTACCCCTTGCTCAGGTTTCAAGAAAGACAGTGAAGAGGCTCATTCGGCATCTTCTTTCTTT 5126
Db 4321 CTACCCCTTGCTCAGGTTTCAAGAAAGACAGTGAAGAGGCTCATTCGGCATCTTCTTTCTTT 4380
Qy 5127 GCAGAAATTTCTCAATATACAAAGTGGCTCCATTTGGGTGTGATATTCCTGGCTGGTATTG 5186
Db 4381 GCAGAAATTTCTCAATATACAAAGTGGCTCCATTTGGGTGTGATATTCCTGGCTGGTATTG 4440
Qy 5187 TGGGTCTCAGTGAAGATGGCATCACCCCTTATCTTGGCTGTGATCTTCTTTTCCAC 5246
Db 4441 TGGGTCTCAGTGAAGATGGCATCACCCCTTATCTTGGCTGTGATCTTCTTTTCCAC 4500
Qy 5247 TATTTACTTGGGGTAACTCCGCCCTGAGGAACTGCATACCAATTTGCAAGAAAGGAGATAC 5306
Db 4501 TATTTACTTGGGGTAACTCCGCCCTGAGGAACTGCATACCAATTTGCAAGAAAGGAGATAC 4560
Qy 5307 AGTGACCTCTGATGCTATCTATCTTTTACCTAATAATTTGTTCTGCTCTTCCAGGAAAT 5366
Db 4561 AGTGACCTCTGATGCTATCTATCTTTTACCTAATAATTTGTTCTGCTCTTCCAGGAAAT 4620
Qy 5367 TGGGATCTGTAAGGCCCTTCTCCAGAGGTGGTGTGAGATCTTCCCTTACTAACTGT 5426
Db 4621 TGGGATCTGTAAGGCCCTTCTCCAGAGGTGGTGTGAGATCTTCCCTTACTAACTGT 4680
Qy 5427 TTGAAGCAAAAAAACCACCGTGTGAGTACCTTAGAAAAAGAAATAGTTTGTATAGAGCTT 5486
Db 4681 TTGAAGCAAAAAAACCACCGTGTGAGTACCTTAGAAAAAGAAATAGTTTGTATAGAGCTT 4740
Qy 5487 CCTGATGATATAGTGTGCTCTCTGAAATCAAGCTTCTCAATTTCAAGTGTGCCCGCTGCA 5546
Db 4741 CCTGATGATATAGTGTGCTCTCTGAAATCAAGCTTCTCAATTTCAAGTGTGCCCGCTGCA 4800
Qy 5547 GATGATGAGCAAAAGCATCTCTCTGCGCTTTTCTGCGGGCTATATGTTCTCAG 5606
Db 4801 GATGATGAGCAAAAGCATCTCTCTGCGCTTTTCTGCGGGCTATATGTTCTCAG 4860

QY 5607 AACATTTGCTGCGGAGAAATTTGTAACGGGGAAGAGTTGGAGCTTGCAATTTTTCACGCA 5666
DB 4861 AACATTTGCTGCGGAGAAATTTGTAACGGGGAAGAGTTGGAGCTTGCAATTTTTCACGCA 4920
QY 5667 CTTCACTGTGAGCGCGAGTCTGCAATTTTCCATAAAATCAGAGAAATCCGAGTGTCTG 5726
DB 4921 CTTCACTGTGAGCGCGAGTCTGCAATTTTCCATAAAATCAGAGAAATCCGAGTGTCTG 4980
QY 5727 GTTGAAGGTAAAGCCAGAGGCTGTGCTATCCAGTCTCTTACTTGAATGAATATCGAGAA 5786
DB 4981 GTTGAAGGTAAAGCCAGAGGCTGTGCTATCCAGTCTCTTACTTGAATGAATATCGAGAA 5040
QY 5787 ACAGACCTTGCCTGAGAGGGGCAACCCCTTCAATTTATCTGAGGGGTATCGGAAG 5846
DB 5041 ACAGACCTTGCCTGAGAGGGGCAACCCCTTCAATTTTCTGAGGGGTATCGGAAG 5098
QY 5847 CTCCTATTGGTCTGCGCAACCACTGCAATTA 5877
DB 5099 CTCCTATTGGTCTGCGCAACCACTGCAATTA 5128

RESULT 5

AR030784 LOCUS AR030784 6395 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5861312.
ACCESSION AR030784
VERSION AR030784.1 GI:5943998
KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

Unclassified.

REFERENCE 1 (bases 1 to 6395)

AUTHORS Varshavsky, A. and Kwon, Y. Tae.

TITLE Nucleic acid encoding mammalian UBR1

JOURNAL Patent: US 5861312-A 1 19-JAN-1999;

FEATURES Location/Qualifiers

source

1..6395

/organism="unknown"

/mol_type="unassigned DNA"

ORIGIN

Query Match 70.3%; Score 4435.4; DB 6; Length 6395;

Best Local Similarity 86.3%; Pred. No. 0;

Matches 4997; Conservative 0; Mismatches 721; Indels 69; Gaps 6;

6;

QY 590 TTCAGGGCGCGTGTAAAGTGTCTGCTCTCTC-TCGACCGGCCACAGGTTCCGCT 648

DB 8 TTCAGGGCGCGTGTAAAGTGTCTGCTCTCTC-TCGACCGGCCACAGGTTCCGCT 67

QY 649 TGCTCTGCGCGGGGTGCGCACTGCGAGCGTCAAGTTTCCCTCAAGATGGCGGACGAGG 708

DB 68 AGCTGGCGCGGGGTGCGCACTGCGAGCGTTCGTTTCCCTTAAGATGGCGGACGAGG 127

QY 709 AGGCTGGAGGTACTGAGAGGATGGAATCAGCGCGGAGTTACCCAGACCCCTCAGCGTC 768

DB 128 AGATGAGCGCGCGGAGAGGATGGAAGTGTCTGCTGCGGTGCGGCGGCCACAGGTTCCGCT 187

QY 769 TGGCATCTTGGTGGGATCAGCAAGTTGATTTTATCTGCTTCTGTCATCATTTGGCAC 828

DB 188 CGGCATCTGTTGGGATCAGCAAGTTGATTTTATCTGCTTCTGTCATCATTTGGCAC 247

QY 829 AATTGGTCCAGAAATTTACTTTGCTGAAATGGACCCAGAGTTGGAAGCAGGAGGAAA 888

DB 248 AATTAGTCCAGAAATTTATTGCTGAGATGGACCCAGAGTTGGAAGCAGGAGGAGA 307

QY 889 GTGTACAAATGTCAATATTCATCTCACTGGAATGTGATTTATTTGGAGAAATCCAGATA 948

DB 308 GTGTACAGATGTCAATATTCATCTCTCTTGGAGTGTGATTTATTTGGAGAGGATCCGATA 367

QY 949 TTTGCTTAGAGAAATTTGAAGCACAGTGGGCAATTTTTCAGCTTTGTTGGAGGCTTTCAAA 1008

DB 368 TTTGCTTAGAGAAATTTAAACACAGTGGAGCGTTCCAGTTGTGTGGGAAAGGTTTCAAAA 427

QY 1009 GTGGAGAGACAACCTATTCTTTCAGGGGATTTGTGCAATTTGATCCAAACATGTGTACTCTGTA 1068
DB 428 GTGGAGAGACAACCTATTCTTTCAGGGGATTTGTGCAATTTGATCCAAACATGTGTACTCTGTA 487
QY 1069 TGGACTGTCTCCAGGACAGTGTTCATAAAATCATCGTTTACAAGATGATCTTCTACTG 1128
DB 488 TGGACTGTCTCCAGGACAGTGTTCATAAAATCATCGTTTACAAGATGATCTTCTACTG 547
QY 1129 GAGGAGGGTCTTGTGACTGTGAGAGACACAGAGGCAATGGAAAATCTGGCCCTTTTGTGTAA 1188
DB 548 GAGGAGGGTCTTGTGACTGTGAGAGACACAGAGGCGTGGAAAATCTGGCCCTTTTGTGTGG 607
QY 1189 ATCATGAACCTGGAGAGGACAGTACTATAAAGAGAAATTCACGCTGTCCGTTGAATGAAG 1248
DB 608 ATCATGAACCTGGAGAGGACAGTACTATAAAGAGAGCTTACATTTGCCCATTTGAATGAAG 667
QY 1249 AGGTAATTTGTCAGAGCCAGGAAATATTTCTTCAGTGTATAAAATATATCTGCTGAGAAAATGA 1308
DB 668 AGGTAATTTGTCAGAGCCAGGAAATATTTCTTCAGTGTATAAAATATATCTGCTGAGAAAATGA 727
QY 1309 CTATATGGGAAGAGGAAANAGAACTGCTCTCTGAACTCCAGATAAGGGAGAAAATAGAA 1368
DB 728 CTATATGGGAAGAGGAAANAGAACTGCTCTCTGAACTCCAGATAAGGGAGAAAATAGAA 787
QY 1369 GATACATTTGTGCTCTTTTCAATGATGAACACCAATTCATATGACCAACATCATATACAGCC 1428
DB 788 GATACATTTGTGCTCTTTTCAACGATGAGCACCAATTCGATATGATCATGTGATCTACAGTC 847
QY 1429 TACAAGAGCTCTTGTGACTGTGAGCTGCGAGAGGCCAGTTCATACCACTGCCATTTGACA 1488
DB 848 TGCAGAGAGCTCTAGATTGCGAGCTTGCAGAGGACACAGCTGCACAGCTGCCATTCGACA 907
QY 1489 AAGAGGTCGTGCGGGCTGTTTAAAGCGGAGCTTATGCTGTTTGCAGGAAGCAAGGAGG 1548
DB 908 AAGAGGTCGTGCGGGCTGTTTAAAGCGGAGCTTATGCTGTTTGCAGGAAGCAAGGAGG 967
QY 1549 ATATAAGAGTCAATTCAGAAAATGTCTCAACATCCACTTCATGTAGTAGATTTACACT 1608
DB 968 ATATAAGAGTCAATTCAGAAAATGTCTCTCAGCACCCCTCCATGTGGAAGTGTGCACT 1027
QY 1609 CAGAGATTATGGCTCATCAGAAAATTTGCTTTGCGTCTTGGTTCTCTGGATGAACAAAATTA 1668
DB 1028 CCGTGGTTATGGCTCATCAGAAAATTTGCTCTGCGCTTGGCTCTCTGGATGAACAAAATTA 1087
QY 1669 TGAGCTATTCAAGTGAATCTTTAGGAGAGTCTTTTCCAGCAATGCTTTAGAGAAACCTG 1728
DB 1088 TGAGCTATTCAAGTGAATCTTTAGAGAGATTTTTCAGGAGCTGCTCGTAGAGAACCTG 1147
QY 1729 ACTCGGAGATCCCTGTCTCTATAAGCAGGTTAAATGCTTTTGGGATGCAAAAGCTTTATAAG 1788
DB 1148 GCTCTGAAAATCCCTGTCTTATAAGCAGACTAAATGCTTTGGGATGCAAAAATTTATAAG 1207
QY 1789 GTGCCGCTNAGATCTTCAATGAATTTGATCTTCAGCAGTCTTTTATGGAGATGGAATACA 1848
DB 1208 GTGCCGCTNAGATCTTCAATGAATTTGATCTTTTAGTAGTCTTTTATGGAGATGGAATACA 1267
QY 1849 AAAAACTCTTTGTATGGAATTTTGAAGTATTAATAACACTGCAGAAAATATATACA 1908
DB 1268 AAAAACTCTTTGTATGGAATTTTGAAGTATTAATAACACTGCAGAAAATATATACA 1327
QY 1909 GTGATGATCATGACAGAGATTTCTATAAATCTGCACTTTTTCAGTTTTCAGATTTTACTGTC 1968
DB 1328 GCGACGACCAACAGAGAGGAGCATCTCCATTAACCGCCCTGTCTGCGTGCAGATGCTCACCGTCC 1387
QY 1969 CTACTCTGGCTCGACATCTTATGAAGAGGAGATTTTCTCTGCTCATTTACTGAAACTC 2028
DB 1388 CGACCTTTGGCCCGCATCTTATGAAGAGGAGATTTTATTTCTGTCTATTTACTGAAACGC 1447
QY 2029 TGCTAGAGTCTTACCTGAGTACTTGGACAGGAAATTAATTCAACTTCAGGAGCTTATA 2088
DB 1448 TGCTAGAGTCTTACCTGAGTACTTGGACAGGAAATTAATTCAACTTCAGGAGCTTATA 1507
QY 2089 GCCAGGACAAATTTGGGAGAGTATATATGACAGTAAATATATGACCTAAAGTATATCTCTGATCA 2148

QY 4300 CTCTGTGCAATCTGTGATCCCATATTTCTTTGCAACCTCAAAAGATATAACAGTGAGA 4359
DB 3728 CTCTGTGCAACATCTGTATCCCATATTTCTTTGCAACCTCAAAAGATATAACAGTGAGA 3787
QY 4360 ATGCAGATGCTCTGTCTCAACTTTTGGACCTGGCACGGTGGATACAGACTGTCTTGGCCA 4419
DB 3788 ATGGCGAGGCTCTGTCTCACTTTTGGACCTGGCACGGTGGATACAGACTGTCTTGGCCA 3847
QY 4420 GAATATCAGGTTTAAATATAAGACATGCTTAAAGGAGAAACC---CAATTCCTATTTTCT 4476
DB 3848 GAATATCAGGTTTAAATATAAGACATGCTTAAAGGAGAAACC---CAATTCCTATTTTCT 3907
QY 4477 TTAATCAGGAAATGGAGATTTCTATTTGGAGTTCCATTCATCTGAGTTTGGCGTTG 4536
DB 3908 TTAATCAGGAAATGGAGATTTCAACTTTTGGAGTTTCAATTCATCTGAGTTTGGAGTTT 3967
QY 4537 AGTCTTCGATTAATAATTAATCAATAGCATCAAGGAAATGTTATTTCTTTTGGCACACAA 4596
DB 3968 AGTCTTCGATTAATAATTAATCAATAGCATCAAGGAAATGTTATTTCTTTTGGCACACAA 4027
QY 4597 TTTATAGAAATGGATTTGAAGTGCACCTGATGAAGGGATCTCGAGTCCCATGCTGA 4656
DB 4028 TTTACAGAAATGGCTTGAAGTGCCTCTGATGAACCTAGACCCACGAGTGCCCATGATGA 4087
QY 4657 CTTGAGACACCTGCGCTTCACTATCCAGGCAATTTGAAATCTATTTGGAGATGAAGGAA 4716
DB 4088 CTTGAGACACCTGCGCTTCACTATCCAGGCAATCGAAACCTGTTTGGAGATGAAGGAA 4147
QY 4717 AACCTCTGTTGGAGCACTTCAAAATAGCAGCATTAATGGTCTGAAACATTAATGCAAT 4776
DB 4148 AACCTCTATTTGGAGCACTTCAAAATAGCAGCATTAATGGTCTGAAAGCGCTTAATGCAAT 4207
QY 4777 TTGCAAGTTCACAGAGGATTAACCTCTCTCAGGCTCTGATACAGAAACATCTGGTTCGTC 4836
DB 4208 TTGCAAGTTCACAGAGGATTAACCTCTCAGGCTCTGATACAGAAACATCTGGTTCGTC 4267
QY 4837 TTCTATCAGTGTCTTCTTAACATAAATCAGAAATACAGAGATACACCATGCTTCTGTCTATAG 4896
DB 4268 TCCTGTCAAGTTATCTTCTCAACCTGCAATCAGAAATACACCATGCTTCTGTCTGTGG 4327
QY 4897 ATCTGTTTCATGTTTGGTGGTGTCTGTAGCATTCCTCATCTTGTATTTGGAGTACC 4956
DB 4328 ATCTCTCCATGTTCTGTGTGGCGCAGTCTTAGCGTTCCCATCTCTGTTATTTGGAGTACA 4387
QY 4957 CTGTGATCTGAGGCTTCTTCAAGTGTCTTCTTATACCAACCTTATCTCTTCCAT 5016
DB 4388 CGTGGATCTGCGGCGTGGCACTTAGTCTTCAATATACCACTCTATCTCTTCCATC 4447
QY 5017 TGATCACCATGCGCACATGCTTCAAGATACATCTATACAGTACAGCA-----CAG 5064
DB 4448 TGATCACCATGCGCACATGCTTCAAGATACATCTATACAGATACAGATCTGTCTCCAG 4507
QY 5065 GCCTACCTCTGCTCAGGTTCAAGAGACAGTGAAGAGCTCATTCGCGATCTTCTTCT 5124
DB 4508 GGCGCCGCTCTGCTGAGGTTGAAGAGTATGTTGAGGAGGCTCGCTGTGCATCTGTCTTCT 4567
QY 5125 TTGCAGAAATTTCTCAATATACAGTGGCTCCATTTGGTGTGATATCTCTGCTGTGAT 5184
DB 4568 TTGTGGAAGTGTGGAGCACACAGCGGCTCACTGGGTGGCGGTCTCCGCGTGGTACC 4627
QY 5185 TGTGGGTCTCACTGAAGAAATGGCATCACCCCTTATCTTCTGCTGTGCTGATTTTCTTCC 5244
DB 4628 TGTGGCTCTCCCTGAGGAAGCGCATCACCCCTTACCTCCGCTGTGCTGCACTGCTTTTCC 4687
QY 5245 ACTATTTACTTGGGTAACTCGGCTCAGGAACTGATACCAATCTTGCAGAGAGAGT 5304
DB 4688 ACTATTTACTTGGAGTAGCTCGGCTGAGAACTGTTTGGCCAACTGTTTGCACACTGCTTTCC 4747
QY 5305 ACAGTGCACTCTGTAGCTATCTATCTTTACCTTACAAATTTGTTCTGCTCTTCCAGGAAT 5364
DB 4748 TCAGTGCACCTCTGTAGCTATCTATCTTTACCCACAAATTTGTTCTGCTCTTCCAGGAAT 4807

QY 5365 ATTGGGATACCTGTAAAGGCGCTTGTCTCAGAGGTGGTGGATCTCTGCTTACTAACT 5424
DB 4808 ATTGGGATACCATTAAGGCCCTTACTACAGAGGTGGTGGATCTCTGCTTACTCAAGT 4867
QY 5425 GTTTGAACCAAAAAACAACCGTGGTCAAGTACCTTAGAAAAAGAAATAGTTTGTATAGAGC 5484
DB 4868 CTTTGAACCAAGAAAGTCTGTGGTCAAGTACCTTAGAAAAAGAAATAGTTTGTATAGAGC 4927
QY 5485 TTCTGTATGATATATAGTGGCTCTCTGATCAAGCTTCTCATTTCAAGTGGCCCAACGGTCTG 5544
DB 4928 TTCTGTAGGACTACAGCTGTCTTCTAAATCAGGCTTCTCACTTTAGGTGTCCACGGTCTG 4987
QY 5545 CAGATGATGAGGAAAGCATCTGTCTCTCTGCTTCTGCTGGGCTATATCTATGTTCTC 5604
DB 4988 CAGATGATGAGGAAAGCATCTGTCTCTCTGCTTCTGCTGGGCTATATCTATGTTCTC 5047
QY 5605 AGAACATTTGCTGCGCAGGAAATTTGAAACGGGAAAGAGTGGAGCTTGGCATTTTTCAGG 5664
DB 5048 AGAACATTTGCTGCGCAGGAAATTTGAAACGGGAAAGAGTGGAGCTTGGAGCTTGGCTTTCATG 5107
QY 5665 CACTTCACTGTGAGCGGAGTCTGCAATTTTCTTAAATCAGAAATGCGAGTGGTCC 5724
DB 5108 CGCTTCAATTTGTGGTGTGGAGTCTGCAATTTTCTTAAATCAGAAATGCGAGGTGGTCC 5167
QY 5725 TGGTTGAAGTAAAGCCAGAGGCTGTGCTCTATCCAGCTCTTACTTGGATGAATAGAG 5784
DB 5168 TGGTTGAAGTAAAGCCAGAGGCTGTGCTCTATCCAGCTCTTACTTGGATGAATAGAG 5227
QY 5785 AAACAGACCTTGGCTGAAGAGGGGCAACCCCTTCTTAAATCAGAAATGCGAGGTGGTCC 5844
DB 5228 AAACAGACCTTGGCTGAAGAGGGGCAACCCCTTCTTAAATCAGAAATGCGAGGTGGTCC 5287
QY 5845 AGCTCCATTTGGTCTGCGCAACCACTGCAATTTATAGAAAGAGATTGCTAGGAGCCAAAGAG 5904
DB 5288 AGCTGCAATTTGGTCTGCGCAACCACTGCAATTTATAGAAAGAGATTGCTAGGAGCCAAAGAG 5347
QY 5905 CTAATCAGATGTTATTTGGATTCAACTGGCAGTTACTGTGAGCTCCAATCTGCTCAG 5964
DB 5348 CTAATCAGATGTTATTTGGATTCAACTGGCAGTTACTGTGAGCTCCAATCTGCTCAG 5407
QY 5965 ACAATCAAAATGACGACAGTAGTAAGAGCTGATTAATAATTTATGAAATCTTCTGAGG 6024
DB 5408 ACAATCAAAATGACGACAGTAGTAAGAGCTGATTAATAATTTATGAAATCTTCTGAGG 5467
QY 6025 GCTGGGAAGTATTTGGAGGCTCTTTGCTCCATGTCCAGGTTCACTTACATCAATAAAT 6084
DB 5468 ACGGGGAAGTATTTGGAGGCTCTTTGATCCATGTCCAGATTCAACACATTAATAAAT 5527
QY 6085 ATTTCTTAATGAGATTTGCTTCAATTTAGCAAAATATGCTTCAAGGAAAAA-AGGAC 6143
DB 5528 ATTTCTTAATGAGATTTGCTTCAATTTAGCAAAATATGCTTCAAGGAAAAA-AGGAC 5587
QY 6144 ATAGATCAATCTGT-----TTT 6160
DB 5588 ATAGATTAATGTTTATGTTCTAGAACACTAAAGAAATGTTGTTTCAATCCAGTGTCTAT 5647
QY 6161 ATGTGCTAGTATTTCCAGGAATTTATTCCTTTCATATTTGCTCTCATTTTCAATTTTATTT 6220
DB 5648 TTCTGCTAATATTTCCAGGAATTTATTCCTTTCATATTTGCTCTCATTTTCAATTTTATTT 5707
QY 6221 CATCACTGTTGATAGAGTCAAGTCAAAAGTGTGTAGACATTTTATGTTGTGTTAAC 6280
DB 5708 CACCACTGTTGATAGAGTCAATTAAGCAATTTGAGCAATTTCTCCATCTGCTAAC 5767
QY 6281 TCTTCTCAATTTTGTATTTTGGTGT 6307
DB 5768 ATCTGTCACCTTTGTATTTTGGTGT 5794

RESULT 6

AR121463

LOCUS

DEFINITION Sequence 1 from patent US 6159732.

6395 bp

DNA

linear

PAT 16-MAY-2001

Db	1868	TAGTGC	AAATTG	TGCGGT	CATAGT	CTGGAA	ACCAAA	TCTCT	CAAAAG	TGCTCT	GAGGAC	CTTGG	1921
Qy	2509	TAAGCATA	CATCTG	CCACTCT	CTAGGAC	CCCTT	GTGTG	TCTT	CATGT	ACGTG	TTTAAG	CAGGC	2568
Db	1928	TAAGCATA	CACCTG	CCACTCT	CTAGACA	CACTT	GCTGT	CTT	CAATG	TACG	TTTAAG	CAAC	1987
Qy	2569	TGGGTG	CTGTTT	CAAGACT	CGATGA	ATTTG	TGTCT	TTT	TGAGG	AGCATTT	CAAGT	AGAGTAC	2628
Db	1988	TAGTGT	GTATTT	CAAGACT	CGATGA	ATTTG	TGTCT	TTT	TGAGG	AGCATTT	CAAGT	AGAGTAC	2047
Qy	2629	TAGTGG	AAATAT	CTCTTT	TACGTT	TGTCT	TGTTG	TCGCC	CAAGT	TGTTG	TCGATG	AGATG	2688
Db	2048	TGTTGG	AGTAT	CCCGCT	CGCTG	CTGCT	TGCTG	CTG	AGGTT	TGTTG	TCGATG	AGATG	2107
Qy	2689	GAAGAA	TGCACT	GTCTCT	TATTAG	CCAGG	TGTTT	TATTA	CCAAAG	TATG	TTTAAG	TGCAGAG	2748
Db	2108	GAAGAA	ACGGCT	CTCACT	CATCAG	CCAGG	TTTCT	TATTA	CAAGAT	TGTTAA	ATGTCAGAG		2167
Qy	2749	AAGAAAT	GTATG	ATAAGAT	CATCAT	CATGCT	TCAGAT	TGTG	CAATCTT	TAATG	GAATCCCA		2808
Db	2168	AGGAAAT	GTACG	ATAAGAT	CATCAT	CTGCT	CAGAT	TGGAG	CACTATA	TAAATG	GAATCCCA		2227
Qy	2809	ATAAGT	CTCTG	TACTG	TACTT	CAGAGG	TATGAA	CTTGC	CGAGG	CTTTTAA	CAAGACCA		2868
Db	2228	ACAAGT	CTCTG	TACTG	TACTT	CAGAGG	TATGAA	CTTGC	CGAGG	CTTTTAA	CAAGACCA		2287
Qy	2869	TATCTA	CAAAAG	ACCGAG	TTGAT	TAAACA	ATATAT	AATACA	CTTAAT	AGAGA	AAATG	CTTC	2928
Db	2288	TATCCA	CAAAAG	ACCGAG	TTGAT	TAAACA	CAATAT	AATACA	CTTAAT	AGAGA	AAATG	CTTC	2347
Qy	2929	AGGTCT	CATCTA	TATTTG	GGGTG	AGCGTT	ATGTG	PACT	CGAGT	GGGAAAT	TGTG	ACCAAG	2988
Db	2348	AGGTCT	CATCTA	TATTTG	GGGAA	CGTTAT	GTG	PACT	CGAGT	GGGAAAT	TGTG	ACCAAG	2407
Qy	2989	AAGAGG	TCACAA	TGAGAG	AAATCA	TTCAC	TTGCT	TTTGC	ATTTGA	ACCCAT	CGCCAC	CAGTG	3048
Db	2408	AGGAGG	TTTAA	TGAGAG	AGATTACT	CAC	TTTACT	TGCA	TTTGA	GGCCAT	CGCCAC	CAGTG	2467
Qy	3049	CCATTG	CCAAAA	TTTACT	CGAGAA	TGAAA	TAATG	AAATG	AAATG	AAATG	AAATG	AAATG	3108
Db	2468	CCATG	CCAGAA	ACCTACT	CTGAGAA	CGAAAT	TAATG	AAATG	AAATG	AAATG	AAATG	AAATG	2527
Qy	3109	ACAAAG	TGGCCA	CATTTA	GAGAA	CCAGGT	GTATG	AGCCCA	TGGAG	TTTATG	AAATG	AAATG	3168
Db	2528	ACAAAG	TGGCCA	CATTTA	GAGAA	CCAGGT	GTATG	AGCCCA	TGGAG	TTTATG	AAATG	AAATG	2587
Qy	3169	ATGAAT	CAC	TGAAG	AGCTTCA	ATATG	CTATTT	TATCA	TCTACT	CAAAA	CCAGCAT	AGCA	3228
Db	2588	ATGAAT	CAC	TGAAG	AGCTTCA	ATATG	CTATTT	TATCA	TCTACT	CAAAA	CCAGCAT	AGCA	2647
Qy	3229	AGGCTG	AAATAT	GTG	CAGAGAA	AGGAG	AAAA	CAAG	AAAA	CAAG	AAAA	CAAG	3288
Db	2648	AGGCTG	AAATAT	GTG	CAGAGAA	AGGAG	AAAA	CAAG	AAAA	CAAG	AAAA	CAAG	2707
Qy	3289	CACCAC	CACTCT	CGAAT	TTCGCC	CTGCTTT	CAGCA	AAAGT	GAATTA	ACCTTT	CTCAACT	GTG	3348
Db	2708	CGCCAC	CTCTCC	AGAGT	TTCTG	CCCTGCTTT	CAGCA	AAAGT	GAATTA	ACCTTT	CTCAACT	GTG	2767
Qy	3349	ATATCAT	GTATG	ATAC	TTCTC	GAGGAC	CGTA	TTT	GAGG	GGGCA	ATPAG	ACATCT	3408
Db	2768	ATGTTAT	GTATG	ATAC	TTCTC	GAGGAC	CGTA	TTT	GAGG	GGGCA	ATPAG	ACATCT	2827
Qy	3409	TGTGG	ACCGA	AGGAT	GTCTT	CCAAAT	TGCTTTT	CATAT	CTG	GCAT	TTTGGG	TTTTACT	3468
Db	2828	TGTGG	ACCGA	AGGAT	GTCTT	CCAAAT	TGCTTTT	CATAT	CTG	GCAT	TTTGGG	TTTTACT	2887
Qy	3469	AGAAG	CAAC	AGCTTT	CAAAA	AGCTCT	CTG	GAAG	AGAG	GTAA	CA	TTTGA	3528
Db	2888	AGAAG	CAG	AGCTTT	CAAAA	AGCTCT	CTG	GAAG	AGAG	GTAA	CA	TTTGA	2947
Qy	3529	CTTCA	AGATTG	GGAAG	TTTCA	GCAT	-----	GAAT	ATACA	AAATG	CTTTT	TGAAAA	3579
Db	2948	CTTCA											

QY 4657 CCTGAGACCTGGCGCTTCTACTATCAGGCAATTAATAATCTATTGGAGATGAAGAA 4716
 Db |||||
 QY 4088 CTTGGAGACAGTGTGGTCTTCCATCCAGGCATCGAATACTGTTGGAGATGAAGAA 4147
 Db |||||
 QY 4717 AACCTCTGTTGGAGACATCTCAAAATAGGACGATTAATGGTCTGAAAGCATTAATGCAAT 4776
 Db |||||
 Db 4148 AACCTCTATTGGAGACATCTCAAAATAGGACGATTAATGGTCTGAAAGCGCTTAATGCAAT 4207
 QY 4777 TTGCAAGTGCACAGAGATTAACCTCTCAGGTCCTGATACAGAAACATCTGGTTCGTC 4836
 Db |||||
 Db 4208 TTGCAAGTGCACAGAGGCTACCTGCCCTCAGTCTGATACAGAAACATCTGGTTCGTC 4267
 QY 4837 TTCTATCAGTGTCTCTCAACATAAATCAGAGATACACATCGCCTTCTGTCATAG 4896
 Db |||||
 Db 4268 TCCCTCAGTATCTCTCAACCTGCAATCAGAAATACACAGCCCTTCTGTCGTGG 4327
 QY 4897 ATCTCTTTTCATGTTTGGTGGTGTGTGTAGCAATCCCATCTCTGATTTGGGATGACC 4956
 Db |||||
 Db 4328 ATCTCTTCCATGTTCTGCTGGCGCAGTCTTAGCGTCCCATCTCTGATTTGGGATGACA 4387
 QY 4957 CTGTTGATCTGAGCCTTCTCAGTGTCTTCTTCTTATTAACCACTTTATCTCTTCCAT 5016
 Db |||||
 Db 4388 CCGTGGATCTGACGCGCTGCGCATCTAGTCTTCTATATAACCACTCTATCTCTTCCATC 4447
 QY 5017 TGATCACCATGCGCACATGCTTCAAGATCTTACAGTACTTACAGTACAGCA - - - - -CAG 5064
 Db |||||
 Db 4448 TGATCACCATGCGCACATGCTTCAAGATCTTACAGTACTTACAGTACAGCACTGCTCCAG 4507
 QY 5065 GCTTACCCCTTCTCAGGTTCAAGAGACAGTGAAGAGGCTCATTTCCGATCTCTTCTTCT 5124
 Db |||||
 Db 4508 GCGCGCGCTTCTGAGGTTGAAGAGATAGTGAAGGCTCGCTGCACTGCTTCTTCT 4567
 QY 5125 TTGCAAGAAATTTCTCAATATACAGTGGCTCAATTTGGGTGTGATTAATCTGCTGCTGAT 5184
 Db |||||
 Db 4568 TTGCGAAGTGTGCGACGACACAGACGCGCTCACTGGGTGCGGTGCTCCCGGCTGGTACC 4627
 QY 5195 TGTGGGTCTCACTGAAGATGACATCACCCCTTATCTTCTGCTGCTGATGATTTTCTTCC 5244
 Db |||||
 Db 4628 TGTGCTCTCTGAGGAAGCGCATCACCCCTTACCTCCGCTGTGCTGCACTGCTTTTCC 4687
 QY 5245 ACTATTTACTTGGGTAACTCCGCTGAGGAATCGATACCAATTTCTGCAAGAGAGAT 5304
 Db |||||
 Db 4688 ACTATTTACTTGGATGTCTCGCTGAGAACTGTTGCCAATTTCTGCTGAAGAGAT 4747
 QY 5305 ACAGTCACTCTGTAGCTATCTTATCTTACCTACAAATTTGTTCTGCTTCCAGGAAT 5364
 Db |||||
 Db 4748 TCAGTGCATCTGTAGCTATCTTATCTTTACCCACAAATTTGTTCTGCTTTTCCAGGAAT 4807
 QY 5365 ATTGGGATCTGTAAGGCCCTTGTCTCCAGAGTGTGTGAGATCCTGCTTACTTAACT 5424
 Db |||||
 Db 4808 ATTGGGATACCAATAAGGCCCTTACTACAGAGTGTGTGAGATCCTGCTTACTCAAGT 4867
 QY 5425 GTTTCGAAGCAAAAAACACCGTGTGAGTACCTAGAACCAAGAAATAGTTTGTATAGAGC 5484
 Db |||||
 Db 4868 CTTTGAAGCAGAAAGTGTGTGCTGAGTACCTAGAACCAAGAAATAGTTTGTATAGAGC 4927
 QY 5485 TTCTGTATGATATAGTCTGCTTCTGATCAATCAAGCTTCTCAATTCAGGTGCGCCAGGCTG 5544
 Db |||||
 Db 4928 TTCTGTAGGACTACAGCTCTCTTCTAAATCAGGCTTCTCACTTTAGGTGCTCCAGGCTCTG 4987
 QY 5545 CAGATGATCAGCAAGCATCTGCTGCTGCTTCTGCTGGGCTATACATGTTCTC 5604
 Db |||||
 Db 4988 CAGATGATGAGCAAGCATCTGCTGCTTCTGCTTCTTCTGTTGGGCTATCTGTTCTC 5047
 QY 5605 AGAACATTTGCTGCCAGGAAATTTGAACCGGGAAGAGGTTGGAGCTTGCATTTTTTTCAGG 5664
 Db |||||
 Db 5048 AGAACATCTGTTGCCAGAAATAGTGAATGGGGAAGGTTGGAGCGTGGTTCATG 5107
 QY 5665 CACTTCACGTGGAGCCGAGTCTGCAATTTCTTAAATCAGAAATCCGAGGTGGTCC 5724
 Db |||||
 Db 5108 CGCTTCATTTGTTGTTGAGTCTGCAATTTCTTAAATCAGAAATCCGAGGTGGTCC 5167
 QY 5725 TGGTTGAAGGTAAAGCAGGCTGTGCTCTATCCAGCTCTTCTTACCTTGGATGAATGGAG 5784

Db 5168 TGGTGAAGAAAGCCAGAGGCTGTGCTACCCAGCCCTTACTTGGATGAATATGAG 5227
 QY 5785 AAACAGACCCCTGGCTGAGAGGGGCAACCCCTTCAATTTATCTCGTGGAGGATCGGA 5844
 Db 5228 AAACAGACCCAGGCTTAAAGAGAGAAACCCACTTCAATTTATCTCGGAGCGGATCGGA 5287
 QY 5845 AGCTCCATTTGGTCTGCAACCAACACTGCATATAGAGAGATTTGCTAGGAGCCAAAGAGA 5904
 Db 5288 AGCTGCATTTGGTCTGGCAACGCACTGCATATAGAGAGATTTGCTCGGAGCCAGAGA 5347
 QY 5905 CTAATCAGATTTATTTGGATTTCAACTGGCAGTTACTGTGAGCTTCAACTCTGCTCAAG 5964
 Db 5348 CTAATCAGATGCTATTTGGATTTAACTGGCACTTACTCTGAGCTTCAAGTCTGCTCAAG 5407
 QY 5965 ACAATCAAAATGACGACAGTAGTAAGGCTGATTCAAAATATAGGAAACCTTTCTGAGG 6024
 Db 5408 ACAATCATGAGTGACATCAATAATAAGACTGATCTAAAATTTCTAGAGAACTTTCTGAGG 5467
 QY 6025 GCTGGGAAAGTATTGGAGGGTCTTTTGTCTCCATGCTCAGGTTCACTTACATCAATAAAT 6084
 Db 5468 ACGGGGAAGTATTGGAGGGTCTTTTGTATCCAGTCCAGATTCACACATTAATAAAT 5527
 QY 6085 ATTTCTTAATAGGATTTGCTTTCAATTTAGCAAAACATATGCTTCAAGGAAAAA - AGGAC 6143
 Db 5528 ATTCCTTAATGGAATATTGCTTTCAATTTATCAAAATAAGCTTCAAGGAAAAAACAAGAC 5587
 QY 6144 ATAGATCAATCTGT - - - - -TTT 6160
 Db 5588 ATAGATTAATGTTTTTATGTTCTAGAACACTAAAGAAATGCTTGTTCATCCAGTGTCTAT 5647
 QY 6161 ATGCTGTAGTATTCCAGGAATTTATCCCTTCAATATTTGCTCAATTTGCTCAATTTTATTT 6220
 Db 5648 TTTGCTTAATTTTCCAGAAACTCTTTCCCTTCAATTAAGCTGCTAGTCAATTTTCAAT 5707
 QY 6221 CATCCACTTGTAGATGAAGTCAAGTCAAGACAGTTGTAGACATTTATGTTGTGTTAAC 6280
 Db 5708 CACCACCTGTTAATGAGGTCAATTAAGCATTTGTGACATTTCTCCATCTGCTAAC 5767
 QY 6281 TCTTCTGCAATTTGTTATTTGTTGTTT 6307
 Db 5768 ATCTGCACTTTGTTATTTGTTGTTT 5794

RESULT 7
 BD245011
 LOCUS Nucleic acid encoding mammalian UB1.
 DEFINITION
 ACCESSION BD245011
 VERSION BD245011.1 GI:33054781
 KEYWORDS JP 2002512771-A/1.
 SOURCE unidentified
 ORGANISM unclassified.
 REFERENCE 1 (bases 1 to 6395)
 AUTHORS Varshavsky, A. and Kwon, Y.T.
 TITLE Nucleic acid encoding mammalian UB1
 JOURNAL Patent: JP 2002512771-A 1 08-MAY-2002;
 CALIFORNIA INSTITUTE OF TECHNOLOGY
 COMMENT OS Unidentified
 PN JP 2002512771-A/1
 PD 08-MAY-2002
 PF 02-DEC-1998 JP 2000523318
 PR 02-DEC-1997 US 08/982956
 PI ALEXANDER VARSHAVSKY, YONG TAE KWON
 PC C12N1/15, A61K31/7088, A61K45/00, A61K48/00, A61P21/00, C07K14/47,
 PC C12N1/21,
 PC C12N5/10, C12N15/09, C12N5/00, C12N15/00
 CC Strandedness: Single;
 CC Topology: Linear;
 CC Nucleic acid encoding mammalian UB1
 FT Key Location/Qualifiers
 FT CDS 115..5385.

Db 4208 TTGAGTTTGCACAGAGGCTACTCCCTCAGGTCTGATACAAAACATCTGGCTCGGC 4267
Qy 4837 TTCTATCAGTTGTTCTTCTTAACATAAATCAGAAAGATACACCATGCTCTTCTGTCTATAG 4896
Db 4268 TCCTGTGAGTTATTCTTCTTAACCTGCAATCAGAAATACACCGGCTCTGTCTGTGG 4327
Qy 4897 ATCTGTTTCATGTTTTGGTGGGTGCTGTGTAGCAATTCCTCATCTTGTATTTGGATGACC 4956
Db 4328 ATCTCTTCATGTTCTGTGTGGCGCAGTCTTAGCGTTTCCCATCTTGTATTTGGATGACA 4387
Qy 4957 GTGTTGATCTGCAGCTCTTCTCAGTTAGTTCTTCTATTAACACCTCTTATCTCTTCCATT 5016
Db 4388 CGTGGATCTGCAGCGCTGCGCCTTAGTTCCTATATAACCACTCTATCTCTTCCATC 4447
Qy 5017 TGATCACCATGGCACACATGCTTTCAGATACATACTTACAGTAGACA -----CAG 5064
Db 4448 TGATCACCATGGCGCACATGCTTTCAGATACCTTTCACACAGATACAGATCTGTCTCCAG 4507
Qy 5065 GCCTACCCCTTGTCTCAGGTTTCAGAAAGACAGTGAAGAGCTCAATTCGGATCTTCTTCT 5124
Db 4508 GCGCCGCTTGTCTGAGGCTGAAGAGGATAGTGAAGAGGCTCGCTGTGCACTCTGCTTCT 4567
Qy 5125 TTGCAGAAATTTCTCAATATACAGTGGCTCCATTTGGGTGTGATATTCTTGGCTGTATT 5184
Db 4568 TTGTGGAAGTGTGCGACACACAGACGCGCTCCTGCGGTGGGTGCTCCCGCTGGTACC 4627
Qy 5185 TGTGGGTCTCACTGAAGAAATGCATCAGCCCTTATCTTCTGCTGTGCTGCAATTTGTTTTCC 5244
Db 4628 TGTGGCTCTCCTGAGGAAGGCACTACCCCTTACTCTCGCTGTGCTGCACTGCTTTTTCC 4687
Qy 5245 ACTATTTACTTGGGGTAACTCGGCTGAGGAACCTGCATACCAATTTCTCGAAGAGAGT 5304
Db 4688 ACTATTTACTTGGAGTAGCTCGCCTGAAGAACTGTTTGCCAAATCTGCTGAAGGAGAT 4747
Qy 5305 ACAGTGCACTGTGAGCTATCTATCTTACCTACAAATTTGTTCTGCTCTTCCAGGAAT 5364
Db 4748 TCAGTGCACTGTGAGCTATCTATCTTACCCACAAATTTGTTCTGCTTTCACAGGAAT 4807
Qy 5365 ATTGGGATCTGTAAAGCCCTTGTCTCCAGAGGTGTGTGAGATCTTCTTCTTAACT 5424
Db 4808 ATTGGGATACCAATAGGCCCTTACTACAGAGGTGTGTGAGATCTTCTTCTTCACTCAAGT 4867
Qy 5425 GTTTGAAGCAAAAACACCGTGTGAGTACCTTACCTAGAAAAGAAATAGTTGTATAGAGC 5484
Db 4868 CTTTGAAGCAGAAAGTGTGTGGTGTGAGTACCTTACCTAGAAAAGAAATAGTTGTATAGAGC 4927
Qy 5485 TTCTGTGACTATAGTGTCTTCTGATCAAGCTTCTCATTTCAAGTGTGCCACCGTCTG 5544
Db 4928 TTCTGTGAGGACTACAGCTGTCTTCTTAAATCAGGCTTCTCACTTTAGGTGTCCACCGTCTG 4987
Qy 5545 CAGATGATGAGCAAGCATCTGTCTCTGCTTCTTCTGTGGGCTATATCTATGTTCTC 5604
Db 4988 CAGATGATGAGCAAGCATCTGTCTCTTCTGTCTTCTGTGGGCTATCTGTGTTCTC 5047
Qy 5605 AGAACAATTGTGCCAGGAAATGTGAAAGGGGAGAGGTTGAGACTTGTCAATTTTTCAG 5664
Db 5048 AGAACAATTGTGCCAGGAAATGTGAAAGGGGAGAGGTTGAGACTTGTGCGGTGTTTTCATG 5107
Qy 5665 CACTTCACTGTGAGCCGAGTCTGCATTTTCTTAAATCAGAAATGCGAGTGTGCTCC 5724
Db 5108 CGCTTCATTTGTGTGCTGAGTCTGCAATTTCTTAAATCAGAAATGCGAGGTTGTGCTC 5167
Qy 5725 TGGTTGAAGGTAAAGCCAGAGGCTGTGCTTATCCAGCTCTTCTTGTGATGAATATGAG 5784
Db 5168 TGGTGAAGGAAAGCCAGAGGCTGTGCTTATCCAGCCCTTCTTGTGATGAATATGAG 5227
Qy 5785 AAACAGACCTTGGCTGAAGAGGGGCAACCCCTTCTTATTTCTGCTGAGCGGTATCGGA 5844
Db 5228 AAACAGACCTTGGCTGAAGAGGAAACCCACTTCTTATTTCTGCGGAGCGGTATCGGA 5287
Qy 5845 AGCTCCATTTGGTCTGGCAACAACTGCATATATAGAGAGATTGCTAGGAGCCAAAGAGA 5904

Db 5288 AGCTGCATTTGGTCTGGCAACAGCACTGCATTTATAGAGAGATTGCTCGGAGCCAGGAGA 5347
Qy 5905 CTAATCAGATGTATTTGGATTCAACTGGCAGTTACTGTGAGCTCCAACTCTGCTCTCAAG 5964
Db 5348 CTAATCAGATGTATTTGGATTCAACTGGCAGTTACTGTGAGCTCTGAGTCTGCTCTCAAG 5407
Qy 5965 ACAAATCAGATGAGTGTAGTAAAGCTGTGATTTCAAAATTTCTAGAGAACTTTCTGAGG 6024
Db 5408 ACAAATCAGATGAGTGTAGTAAAGCTGTGATTTCAAAATTTCTAGAGAACTTTCTGAGG 5467
Qy 6025 GCTGGAAAGTATTTGGAGGGTCTTTTGTCCATGTCAGGTTCACTTACATCAATAAAT 6084
Db 5468 ACGGGGGAGTATTTGGAGGGTCTTTTGTATCAATGTCAGATTCACACATTAATAAAT 5527
Qy 6085 ATTTCTTAATGGAGTATTTGCTTCAATTTAGCAACATATGTTTACAGGAAAAA-AGGAC 6143
Db 5528 ATTTCTTAATGGATATTTGCTTCAATTTATCAACATAAGCTTCAAGGGAAAAACAAGAC 5587
Qy 6144 ATAGATCAATCTGT-----TTT 6160
Db 5588 ATAGATTAATGTTTATGTTCTAGAACACTAAAGAAATGCTTGTTCATCCAGTGTCTAT 5647
Qy 6161 ATGCTGTAGTATTTCCAGGAATTTATCCCTTCTATTAATTTGTCTCATTTTCAATTTTATTT 6220
Db 5648 TTCTGCTAATATTTCCAGAAACTCTTCCCTTCATTAACCTGCTAGTTCATTTCTATAT 5707
Qy 6221 CATCCACTGTGTAGTGAAGTCACTCAACAGTTGTGTAGACATTTTATGTGTGTTAAAC 6280
Db 5708 CACCACCTGTGTTAATGAGGTCACTTAAGCATTTGTGGACATTTCTCCATCTGSCTAAC 5767
Qy 6281 TCCTTCTGCAATTTGTTATTTGCTGTTT 6307
Db 5768 ATCTCTGCACCTTGTATTTGCTGTTT 5794

RESULT 8
AF061555 6395 bp mRNA linear ROD 10-JUL-1998
LOCUS Mus musculus ubiquitin-protein ligase E3-alpha (Ubr1) mRNA,
DEFINITION complete cds.
ACCESSION AF061555
VERSION AF061555.1 GI:3170886
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 6395)
AUTHORS Kwon,Y.T. and Varshavsky,A.
TITLE Direct Submission
JOURNAL Submitted (24-APR-1998) Division of Biology, 147-75, California
Institute of Technology, 1200 E. California Bl., Pasadena, CA
91125, USA
FEATURES
source
1. .6395
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/chromosome="2"
/cell_line="MEL-C19"
1. .6395
/gene="Ubr1"
1. .114
/gene="Ubr1"
115. .5388

CQ728565 LOCUS CQ728565 4529 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 14499 from Patent WO02068579.
ACCESSION CQ728565
VERSION CQ728565.1 GI:42297720
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 14499 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
Location/Qualifiers
Source 1..4529
ORIGIN
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
Query Match 55.1%; Score 3473.6; DB 6; Length 4529;
Best Local Similarity 82.5%; Pred. No. 0;
Matches 4475; Conservative 0; Mismatches 24; Indels 927; Gaps 6;
QY 550 CGGGCCGGAAGCACTCTCGAGTCGCTGCTCAAAACCGACCTTCAGGGCCGTCGTAAG 609
DB 1 CGGGCCGGAAGCACTCTCGAGTCGCTGCTCAAAACCGACCTTCAGGGCCGTCGTAAG 60
QY 610 TGTGCTGCTCTCTCTCGACCGGCCACAGGTTTCGCTTCGCTTCGCGCGGGGTCGGC 669
DB 61 TGTGCTGCTCTCTCTCGACCGGCCACAGGTTTCGCTTCGCTTCGCGCGGGGTCGGC 120
QY 670 AACTGCGGCGTCAGTTTCCTCAAGATGGCGGAGAGGCTTCGAGGATACAGAGGA 729
DB 121 AACTGCGGCGTCAGTTTCCTCAAGATGGCGGAGAGGCTTCGAGGATACAGAGGA 180
QY 730 TGGAAATCAGCGCGAGTTACCCAGACCCCTCAGCGTCGCGATCTTGGTGGATCAGC 789
DB 181 TGGAAATCAGCGCGAGTTACCCAGACCCCTCAGCGTCGCGATCTTGGTGGATCAGC 240
QY 790 AAGTTGATTTTATATCTGCTTTCTTGCACTATTTGGCACAATTTGGTCCAGAAATTTACT 849
DB 241 AAGTTGATTTTATATCTGCTTTCTTGCACTATTTGGCACAATTTGGTCCAGAAATTTACT 300
QY 850 TTGCTGAAATGGACCCAGACTTGGAAAGCAGGAGAAAGTGTAACAATGTCAATATTC 909
DB 301 TTGCTGAAATGGACCCAGACTTGGAAAGCAGGAGAAAGTGTAACAATGTCAATATTC 360
QY 910 CTCCACTGGATGGTACTTATTTGGAGAGATCCAGATATTTGCTTAGAGAAATTTGAGC 969
DB 361 CTCCACTGGATGGTACTTATTTGGAGAGATCCAGATATTTGCTTAGAGAAATTTGAGC 420
QY 970 ACAGTGGAGCAATTCAGCTTTGTGGAGGGTTTTCAAAAGTGGAGAGCAACCTATTCTT 1029
DB 421 ACAGTGGAGCAATTCAGCTTTGTGGAGGGTTTTCAAAAGTGGAGAGCAACCTATTCTT 480
QY 1030 GCAGGATTTGCAATTTGATCCAAATGTGACTCTGTATGAGCTGCTTCAGGACAGTG 1089
DB 481 GCAGGATTTGCAATTTGATCCAAATGTGACTCTGTATGAGCTGCTTCAGGACAGTG 540
QY 1090 TTCAATAAATCATCGTTTACAGATGCATCTTCTACTGGAGGAGGGTTCTGTGACTGTG 1149
DB 541 TTCAATAAATCATCGTTTACAGATGCATCTTCTACTGGAGGAGGGTTCTGTGACTGTG 600
QY 1150 GAGACACAGGCGATGGAATACTGGCCCTTTTGTGTAATCATGAACTCTGGAAGAGCAG 1209
DB 601 GAGACACAGGCGATGGAATACTGGCCCTTTTGTGTAATCATGAACTCTGGAAGAGCAG 660
QY 1210 GTACTATAAAGAGAAATTCACGCTGTCCGTTTGAATGAAGAGGTAAATTTGTCAGGCCAGGA 1269

DB 661 GTACTATAAAGAGAAATTCACGCTGTCCGTTTGAATGAAGAGGTAAATTTCTCAAGCCAGA 720
QY 1270 AAATATTTTCTTTCAGTGATATAATATGTCGTAGAAATGACTATATGGGAAGGAAAAAG 1329
DB 721 AAATATTTTCTTTCAGTGATATAATATGTCGTAGAAATGACTATATGGGAAGGAAAAAG 780
QY 1330 AACTGCTCTCTGAACTCCAGATAAGGGAGAAAAAATGAAAGATATATTTGTGTCTTTTCA 1389
DB 781 AACTGCTCTCTGAACTCCAGATAAGGGAGAAAAAATGAAAGATATATTTGTGTCTTTTCA 840
QY 1390 ATGATGAACACCATTCATATGACCACTCATATACAGCTCATACAGCTTCAAGAGCTCTTGACTGTG 1449
DB 841 ATGATGAACACCATTCATATGACCACTCATATACAGCTTCAAGAGCTCTTGACTGTG 900
QY 1450 AGCTCGCAGAGGCCAGTTGCATACCACTGCTCCATTTGACAAAGAGGGTCTCGGGCTGTTA 1509
DB 901 AGCTCGCAGAGGCCAGTTGCATACCACTGCTCCATTTGACAAAGAGGGTCTCGGGCTGTTA 960
QY 1510 AAGCGGAGCTTATGCTGCTTGCAGGAAGCAAAAGGAAGATATAAAGAGTCATTCAGAAA 1569
DB 961 AAGCGGAGCTTATGCTGCTTGCAGGAAGCAAAAGGAAGATATAAAGAGTCATTCAGAAA 1020
QY 1570 ATGTCTCTCAACATCCACTTCATGTAGAGATTTACACTCAGAGATTTATGCTCATCAGA 1629
DB 1021 ATGTCTCTCAACATCCACTTCATGTAGAGATTTACACTCAGAGATTTATGCTCATCAGA 1080
QY 1630 AATTGCTTTGCGTCTTGGTTCCTGGATGAACAAAATTTATGAGCTATTTCAAGTGAATTTA 1689
DB 1081 AATTGCTTTGCGTCTTGGTTCCTGGATGAACAAAATTTATGAGCTATTTCAAGTGAATTTA 1140
QY 1690 GGCAAGATCTTTTTCAGAGCAATGCTTACAGAGAAACCTGACTCGGAGAAATCCCTGTCTCA 1749
DB 1141 GGCAAGATCTTTTTCAGAGCAATGCTTACAGAGAAACCTGACTCGGAGAAATCCCTGTCTCA 1200
QY 1750 TAACAGGTTAACTGCTTGGATGCAAGCTTTTATAAAGTTCGCGTGAAGATCCTTCTCATG 1809
DB 1201 TAACAGGTTAACTGCTTGGATGCAAGCTTTTATAAAGTTCGCGTGAAGATCCTTCTCATG 1260
QY 1810 AATTGATCTTCAGCAGATTTTATGAGATGGAATACAAAACCTCTTTGCTATGGAAT 1869
DB 1261 AATTGATCTTCAGCAGATTTTATGAGATGGAATACAAAACCTCTTTGCTATGGAAT 1320
QY 1870 TTGTGAAGTATTAATAACACTCGCAGAAAGATATATCAGTGATGATCATCAGAGAGTA 1929
DB 1321 TTGTGAAGTATTAATAACACTCGCAGAAAGATATATCAGTGATGATCATCAGAGAGTA 1380
QY 1930 TCTCTATACTGCACTTTTCACTGATGTTTACCTGCTTCTACTCTGCTCGACATCTTA 1989
DB 1381 TCTCTATACTGCACTTTTCACTGATGTTTACCTGCTTCTACTCTGCTCGACATCTTA 1440
QY 1990 TTGAAGAGCAGAAATGTTTATCTGCTCAATCTGTAACCTCTGCTAGAGTCTTCTGAGT 2049
DB 1441 TTGAAGAGCAGAAATGTTTATCTGCTCAATCTGTAACCTCTGCTAGAGTCTTCTGAGT 1500
QY 2050 ACTTGGCAGGAAACATAAATTTCAACTTCCAGGGTTATAGCCAGGACAAATTTGGGAAGAG 2109
DB 1501 ACTTGGCAGGAAACATAAATTTCAACTTCCAGGGTTATAGCCAGGACAAATTTGGGAAGAG 1560
QY 2110 TATATGAGTAAATATGTAACCTAAAGTATATCTCTGATCAGCAAAACCCCAATATGGAAG 2169
DB 1561 TATATGAGTAAATATGTAACCTAAAGTATATCTCTGATCAGCAAAACCCCAATATGGAAG 1620
QY 2170 AAAGATTAAGAAATGCAAGTCTTCCAGGTTTCCAGTCTTTTTCAGAGATCTTACCTCTGA 2229
DB 1621 AAAGATTAAGAAATGCAAGTCTTCCAGGTTTCCAGTCTTTTTCAGAGATCTTACCTCTGA 1680
QY 2230 TGCAAGGAAATGGAAGAAATCCGAGACAGGTTGGGCAACACATTTGAAGTGGATCCTGAT 2289
DB 1681 TGCAAGGAAATGGAAGAAATCCGAGACAGGTTGGGCAACACATTTGAAGTGGATCCTGAT 1740
QY 2290 GGGAGGCTGCAATTTGCTATACAGATGCAATTTGAAGATATTTTATCTCATGTTTCAAGAT 2349

Db 1741 GGGAGGCTGCCATTGCTATACAGATGCAATTTGAAGAAATATTTTACTCATGTTCACAGAGT 1800
Qy 2350 GGTGTGCTTGATGAGAACTCTTACTTGTGGCTTATAAGAAATGTACACAAAGCTGTGA 2409
Db 1801 GGTGTGCTTGATGAGAACTCTTACTTGTGGCTTATAAGAAATGTACACAAAGCTGTGA 1860
Qy 2410 TGAGGTGAGTACCAAGTTTTCATATCTAGTAGCAAGACAGTAGTACAATCGTGTGGACATA 2469
Db 1861 TGAGGTGAGTACCAAGTTTTCATATCTAGTAGCAAGACAGTAGTACAATCGTGTGGACATA 1920
Qy 2470 GTTTGGAAACAAAGCTTCAAGATCTGAGGATCTTTGTAAGCATACATCTGCCACTCT 2529
Db 1921 GTTTGGAAACAAAGCTTCAAGATCTGAGGATCTTTGTAAGCATACATCTGCCACTCT 1980
Qy 2530 CTAGGACCTTGTGGCTTTCATGTACGTTTAAAGCAGGCTGGGTCTGTTTCAAGACTGC 2589
Db 1981 CTAGGACCTTGTGGCTTTCATGTACGTTTAAAGCAGGCTGGGTCTGTTTCAAGACTGC 2040
Qy 2590 ATGAATTTGTCTTTTTCAGGACTTTCAAGTAGAGGTACTAGTGGAAATATCTTTACGTT 2649
Db 2041 ATGAATTTGTCTTTTTCAGGACTTTCAAGTAGAGGTACTAGTGGAAATATCTTTACGTT 2100
Qy 2650 GTCTGGTGTGGTTCGCCAGGTGTGTCTGAGATGTGGCGAAGAAATGGACTGTCTCTTA 2709
Db 2101 GTCTGGTGTGGTTCGCCAGGTGTGTCTGAGATGTGGCGAAGAAATGGACTGTCTCTTA 2160
Qy 2710 TTAGCCAGGTGTTTATTTACCAAGATGTTAAGTGCAGAGAAATGTATGATAAGATA 2769
Db 2161 TTAGCCAGGTGTTTATTTACCAAGATGTTAAGTGCAGAGAAATGTATGATAAGATA 2220
Qy 2770 TCATCATGCTT-----CAGATTGGTGCATCTTTAA 2799
Db 2221 TCATCATGCTTCAGAGAAATATGATGCTGGATTTCTTTCAAGATTGGTGCATCTTTAA 2280
Qy 2800 TGGATCCCAATAAGTTCTTTGTACTGTGTACTTCAGAGGTATGAACCTTGCAGGCTTTTA 2859
Db 2281 TGGATCCCAATAAGTTCTTTGTACTGTGTACTTCAGAGGTATGAACCTTGCAGGCTTTTA 2340
Qy 2860 ACAAGACCATATCTACAAAGACCGAGATTGATTTAAACAATATATACACTAATAGAA 2919
Db 2341 ACAAGACCATATCTACAAAGACCGAGATTGATTTAAACAATATATACACTAATAGAA 2400
Qy 2920 AAATGCTTCAGTCTCATCTATATTTGTGGGTGAGCGTTATGTACTCGAGTGGGAATG 2979
Db 2401 AAATGCTTCAGTCTCATCTATATTTG----- 2429
Qy 2980 TGACCAAGAGAGGTCAAAATGAGAGAAATCAATTCATCTGCTTTGCAATGAACCCATGC 3039
Db 2430 ----- 2429
Qy 3040 CACACAGTGCCATTGCCAAAATTTTACCTGAGAAATGAATAATGAACCTGGCTTAGAGA 3099
Db 2430 ----- 2429
Qy 3100 ATGTCATAAACAAAGTGGCCACATTTAAGAAACAGGTGTATCAGGCCATGAGTTTATG 3159
Db 2430 ----- 2429
Qy 3160 AACTAAAGATGAATCACTGAAGACTTCAATATGTAATCTTTTATCAATTAATCCAAACCC 3219
Db 2430 ----- 2429
Qy 3220 AGCATAGCAAGGCTGAACATATGCAAGAAAGAGAGAAACAGAAACAAAGATGAAG 3279
Db 2430 -----G 2430
Qy 3280 CATTCGCCACCAACCACTCTGAAATTCCTGCCCTGCTTTTCAGCAAGAGTATTAACCTTC 3339
Db 2431 CATTCGCCACCAACCACTCTGAAATTCCTGCCCTGCTTTTCAGCAAGAGTATTAACCTTC 2490
Qy 3340 TCAACTGTGATATCATATGATGATCAATCTCAGGACCGTATTTGAGCGGGCAATAGACAG 3399
Db 2491 TCAACTGTGATATCATATGATGATCAATCTCAGGACCGTATTTGAGCGGGCAATAGACAG 2550

Qy 3400 ATTCTAACTTGTGGACCGAAGGATGCTCCAAATGGCTTTTCATATTTCTGCAATGGGTT 3459
Db 2551 ATTCTAACTTGTGGACCGAAGGATGCTCCAAATGGCTTTTCATATTTCTGCAATGGGTT 2610
Qy 3460 TACTAGAAGAGAACAAACAGCTTCAAAAGCTCCTGAAGAGAGTAACTTTGACATTTT 3519
Db 2611 TACTAGAAGAGAACAAACAGCTTCAAAAGCTCCTGAAGAGAGTAACTTTGACATTTT 2670
Qy 3520 ATCATAGGCTTCAAGATTGGGAAGTTTCAAGCCTAGATATACAAATGCTTTTGGAAAC 3579
Db 2671 ATCATAGGCTTCAAGATTGGGAAGTTTCAAGCCTAGATATACAAATGCTTTTGGAAAC 2730
Qy 3580 TCAAGGAATTTCCCAAGTTAAGAGGCCAGAGGACATGATAAAGCTGATATCTTCAGATGT 3639
Db 2731 TCAAGGAATTTCCCAAGTTAAGAGGCCAGAGGACATGATAAAGCTGATATCTTCAGATGT 2790
Qy 3640 TTGACACAGTGAAGCGATTAAAGAGAAAATCTTGTGTTAAATTTGTAGCAACCAATCAGAT 3699
Db 2791 TTGACACAGTGAAGCGATTAAAGAGAAAATCTTGTGTTAAATTTGTAGCAACCAATCAGAT 2850
Qy 3700 CGGAATCTATTAAGAAATGATGAGATTACTCATGATAAGAAAGCAGAGCAAGAAAGAA 3759
Db 2851 CGGAATCTATTAAGAAATGATGAGATTACTCATGATAAGAAAGCAGAGCAAGAAAGAA 2910
Qy 3760 AAGCTGAAGCTGTAGGCTTACATCGCAGAGATCATGGCTCAGATGTCTGCCCTTACAGA 3819
Db 2911 AAGCTGAAGCTGTAGGCTTACATCGCAGAGATCATGGCTCAGATGTCTGCCCTTACAGA 2970
Qy 3820 AAAATCTTCAATGAACCTATAAATCTATGATGATGATAATCATCAAGAAATGCTCGGAAAG 3879
Db 2971 AAAATCTTCAATGAACCTATAAATCTATGATGATGATAATCATCAAGAAATGCTCGGAAAG 3030
Qy 3880 AAGATTCCATTTAGAGAAAGAGAGACCCAGCAGTCAGTGACTACTCTAGAAATGCTT 3939
Db 3031 AAGATTCCATTTAGAGAAAGAGAGACCCAGCAGTCAGTGACTACTCTAGAAATGCTT 3090
Qy 3940 TGGTCTCTAAACGGGTCCATCTGTTACTGAAAAGAGGTGCTGACGTGCAATCTTTGCC 3999
Db 3091 TGGTCTCTAAACGGGTCCATCTGTTACTGAAAAGAGGTGCTGACGTGCAATCTTTGCC 3150
Qy 4000 AAGAAAGACAGGAGTGAATAATAGAAATATGCAATGATGATAATGCGGCTGTGTCCAGA 4059
Db 3151 AAGAAAGACAGGAGTGAATAATAGAAATATGCAATGATGATAATGCGGCTGTGTCCAGA 3210
Qy 4060 AATCTACTGCTTAAACCCAGCAGGGGAAACCCATAGACTCTCAGGAGAGCCCTAG 4119
Db 3211 AATCTACTGCTTAAACCCAGCAGGGGAAACCCATAGACTCTCAGGAGAGCCCTAG 3270
Qy 4120 ACCCACTTTTTCATGATTCAGACTTTGGCATATGAACTTATACAGGAAGCTGTGTCTCATG 4179
Db 3271 ACCCACTTTTTCATGATTCAGACTTTGGCATATGAACTTATACAGGAAGCTGTGTCTCATG 3330
Qy 4180 TAATGACGCAAGTGTGTCGAGAGTATTTTGAAGCTGTATACAGTCAAGCTCTCAGCAGC 4239
Db 3331 TAATGACGCAAGTGTGTCGAGAGTATTTTGAAGCTGTATACAGTCAAGCTCTCAGCAGC 3390
Qy 4240 GCATTCAATGTTGACTTTTGAATTTGAAAGTGAAGATATCTTTGCGCTCTTTGCAAT 4299
Db 3391 GCATTCAATGTTGACTTTTGAATTTGAAAGTGAAGATATCTTTGCGCTCTTTGCAAT 3450
Qy 4300 CTCTGTGCAATGATGATCCCAATTTCTTTGCAACCTCAAAAGATAAAGTGAAG 4359
Db 3451 CTCTGTGCAATGATGATCCCAATTTCTTTGCAACCTCAAAAGATAAAGTGAAG----- 3505
Qy 4360 ATGAGATGCTCTTTGCTCAACTTTTGAACCCCTGGCAAGTGGATACAGACTGTTCTGGCCA 4419
Db 3506 ----- 3505
Qy 4420 GAATATCAGGTTATATATAGACATGCTAAAGAGAGAAACCCCAATTCCTATTTCTTTA 4479
Db 3506 ----- 3505


```
QY 4480 ATCAAGGAATGGAGATTTCTACTTTGGAGTTCCATTCCATCTGAGTTTGGCGTTGAGT 4539
Db 3506 ----- 3505
QY 4540 CTTGATTAATAATTAATCAATAGCAATCAAGGAATGTTTATTTCTCTTTGCCCAACAATTT 4599
Db 3506 ---GATTAAATAATTAATAGCAATCAAGGAATGTTTATTTCTCTTTGCCCAACAATTT 3561
QY 4600 ATAGAAATTTGAATGGAATGCACTGATGAAAGGATCTCGAGTCCCAATCTGACCT 4659
Db 3562 ATAGAAATTTGAATGGAATGCACTGATGAAAGGATCTCGAGTCCCAATCTGACCT 3621
QY 4660 GGAGCAGCTGGCTTTCACTATCCAGGCAATTTGAAATCTATTGGGAGATGAAGGAAC 4719
Db 3622 GGAGCAGCTGGCTTTCACTATCCAGGCAATTT-----GATAAG 3660
QY 4720 CTCTGTTTGGAGCACTTCAAAATAGGAGCAATAATGGTCTGAAAGCAATAATGCAAGTTG 4779
Db 3661 CTCTAAGAAACTGGTTTGGAGGTTTCAATAATGGTCTGAAAGCAATAATGCAAGTTG 3720
QY 4780 CAGTTGCAAGAGGATTAATCTGCTCAGGTCCTGATACAGAAACATCTGGTTGCTCTTC 4839
Db 3721 CAGTTGCAAGAGGATTAATCTGCTCAGGTCCTGATACAGAAACATCTGGTTGCTCTTC 3780
QY 4840 TATCAGTTGTTCTTCCCTAACATTAATCAAGAGATACACCATGCTTCTGCTATAGATC 4899
Db 3781 TATCAGTTGTTCTTCCCTAACATTAATCAAGAGATACACCATGCTTCTGCTATAGATC 3840
QY 4900 TGTTTCATGTTTGGTGGGCTGTTGTTAGCAATCCCATCTGTTATTGGGATGACCCCTG 4959
Db 3841 TGTTTCATGTTTGGTGGGCTGTTGTTAGCAATCCCATCTGTTATTGGGATGACCCCTG 3900
QY 4960 TTGATCTGCAGCCTTCTTCAGTTAGTTCTTCTATAACCACTTTTATCTTCCATTTGA 5019
Db 3901 TTGATCTGCAGCCTTCTTCAGTTAGTTCTTCTATAACCACTTTTATCTTCCATTTGA 3960
QY 5020 TCACATGGGACACATGCTTACAGTACTACTTACAGTACAGACAGGCTACCCCTTGCTC 5079
Db 3961 TCACATGGGACACATGCTTACAGTACTACTTACAGTACAGACAG----- 4005
QY 5080 AGGTTCAAGAGACAGTGAAGAGGCTCATTCCGCACTCTTCTTTTGCAGAAATTTCTC 5139
Db 4006 ----- 4005
QY 5140 AATATACAAGTGGCTCCATTGGGTGTGATATTCCTGGCTGATATTTGGGCTCACTGA 5199
Db 4006 ----- 4005
QY 5200 AGAATGGCATCACCCCTTATCTTGCTGTGCTGCATTTGTTTCCACTATTTACTTGGG 5259
Db 4006 ----- 4005
QY 5260 TAACTCCGCTGAGGAACATGATACCAATCTGCAGAGGAGGTACAGTGCACCTCTGTA 5319
Db 4006 -----ATTCTGCAAGAGGAGGTACAGTGCACCTCTGTA 4038
QY 5320 GCTATCTATCTTTACCAAAATTTGTTCTGCTCTCTCCAGGAATATGGGATACTGTAA 5379
Db 4039 GCTATCTATCTTTACCAAAATTTGTTCTGCTCTCTCCAGGAATATGGGATACTGTAA 4098
QY 5380 GGCCCTTGCTCCAGAGGTGGTGCAGATCCTGCTTACTAACTGTTTGAAGCAAAAAA 5439
Db 4099 GGCCCTTGCTCCAG----- 4112
QY 5440 ACACCGTGGTCAAGTACCTTAGAAAAAGAAATAGTTTGTATAGAGCTTCTGATGACTATA 5499
Db 4113 ----- 4112
QY 5500 GCTGCTCTCTGAATCAAGCTTCTCATTTCAAGTGGCCACGCTCTGCAGATGATGAGGAA 5559
Db 4113 -----AGTGGCCACGCTCTGCAGATGATGAGGAA 4143
QY 5560 AGCATCTGCTCTGCTGCTTTTCTGTGGGGCTATATCTATGTTCTCAGAAACATTTGCTGCC 5619
```

```
Db 4144 AGCATCTGCTCTGCTGCTTTTCTGTGGGGCTATCTATGTTCTCAGAACATTTGCTGCC 4203
QY 5620 AGGAAATTTGTAACGGGGAAGAGGTTGGAGCTTGCATTTTTCACGCACTTCACTGTGGAG 5679
Db 4204 AGGAAATTTGTAACGGGGAAGAGGTTGGAGCTTGCATTTTTCACGCACTTCACTGTGGAG 4263
QY 5680 CCGAGTCTGCATTTTCTTAAATTCAGAGAAATGCCGAGTGGTCTGTTGAAGTAAAG 5739
Db 4264 CCGAGTCTGCATTTTCTTAAATTCAGAGAAATGCCGAGTGGTCTGTTGAAGTAAAG 4323
QY 5740 CCAGAGGCTGCTTATCCAGCTCTTACTTGGATGAATATGGAGAAACAGACCTTGCC 5799
Db 4324 CCAGAGGCTGCTTATCCAGCTCTTACTTGGATGAATATGGAGAAACAGACCTTGCC 4383
QY 5800 TGAAGAGGGGCAACCCCTTCAATTTATCTCGTAGCGGTATCGGAAGCTCCATTTGGTCT 5859
Db 4384 TGAAGAGGGGCAACCCCTTCAATTTATCTCGTAGCGGTATCGGAAGCTCCATTTGGTCT 4443
QY 5860 GGCAACACACTGCATTTATAGAGAGATTCCTAGAGCCAGAGACTAATCAGATGTTAT 5919
Db 4444 GGCAACACACTGCATTTATAGAGAGATTCCTAGAGCCAGAGACTAATCAGATGTTAT 4503
QY 5920 TTGCAATCAACTGGCAGTTACTGTGA 5945
Db 4504 TTGCAATCAACTGGCAGTTACTGTGA 4529

RESULT 10
AX714232 LOCUS AX714232 3059 bp DNA linear PAT 15-APR-2003
DEFINITION Sequence 916 from Patent BP1293569.
ACCESSION AX714232
VERSION AX714232.1 GI:29889184
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Isegai, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S.,
Yamamoto, J.I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R.,
Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahara, K. and
Masuho, Y.
TITLE Full-length cDNAs
JOURNAL Patent: EP 1293569-A 916 19-MAR-2003;
Helix Research Institute (JP) ; Research Association for
Biotechnology (JP)
FEATURES
source location/Qualifiers
1. 3059
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 48.4%; Score 3055.8; DB 6; Length 3059;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3057; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 680 GTGATTTTCCCTCAAGATGCGGACGAGGCTTGGAGGTACTGAGAGATGGAATCAG 739
Db 1 GTGATTTTCCCTCAAGATGCGGACGAGGCTTGGAGGTACTGAGAGATGGAATCAG 60
QY 740 CGCGAGTTACCCAGACCCCTCAGCGTCTGGCATCTTGGGATCAGCAAGTTGATTT 799
Db 61 CGCGAGTTACCCAGACCCCTCAGCGTCTGGCATCTTGGGATCAGCAAGTTGATTT 120
QY 800 TTATACCTGCTTTCTTGCATCATTTGGCACAATTTGGTGCAGAAATTTTACTTGTGTAAT 859
Db 121 TTATACCTGCTTTCTTGCATCATTTGGCACAATTTGGTGCAGAAATTTTACTTGTGTAAT 180
QY 860 GGACCCAGACTTGAAAGACGAGGAAAGTGTACAAATGTCAATATTTCACTCCACTGGA 919
```

Db 181 GGACCAGACTTGGAAGAGCAGGAGGAAGGTGTACAAATGTCAAAATTTCACTCCACTGGA 240
Qy 920 ATGGTACTTATTTGGAGAGATCCAGATATTTGGCTTAGAGAAATGAGCACAGTGGAGC 979
Db 241 ATGGTACTTATTTGGAGAGATCCAGATATTTGGCTTAGAGAAATGAGCACAGTGGAGC 300
Qy 980 ATTTTCAGCTTTGTGGAGGGTTTTCAAAAGTGGAGAGACCACTTATTTCTTCGACGGATTTG 1039
Db 301 ATTTTCAGCTTTGTGGAGGGTTTTCAAAAGTGGAGAGACCACTTATTTCTTCGACGGATTTG 360
Qy 1040 TGCATTTGATCCAACTGATGTACTCTGTATGGAATGCTTCCAGGACAGTGTTCATATAAAA 1099
Db 361 TGCATTTGATCCAACTGATGTACTCTGTATGGAATGCTTCCAGGACAGTGTTCATATAAAA 420
Qy 1100 TCATCGTTACAGATGCATCTTCTATCTGGAGGAGGGTTCTGTGACTGTGGAGACACAGA 1159
Db 421 TCATCGTTACAGATGCATCTTCTATCTGGAGGAGGGTTCTGTGACTGTGGAGACACAGA 480
Qy 1160 GGCAATGGAAGAACTGGCCCTTTTGTGTAAATCATGAACTTGGAAAGCAGGTACTATAAAA 1219
Db 481 GGCAATGGAAGAACTGGCCCTTTTGTGTAAATCATGAACTTGGAAAGCAGGTACTATAAAA 540
Qy 1220 AGAGAAATTCACGCTGTCCGTTGAAATGAAGAGGTAAATTTGTCCAGCCAGGAAAAATTTTCC 1279
Db 541 AGAGAAATTCACGCTGTCCGTTGAAATGAAGAGGTAAATTTGTCCAGCCAGGAAAAATTTTCC 600
Qy 1280 TTCAAGTATAAATATGTCTGTAGAAATGCACTATATATGGGAAGAGGAAAAAGAACTGCCCTCC 1339
Db 601 TTCAAGTATAAATATGTCTGTAGAAATGCACTATATATGGGAAGAGGAAAAAGAACTGCCCTCC 660
Qy 1340 TGAATCCAGATATAGGGAGAAAAATGAAAGATATCTATTGTGTCTTTTCAATGATGAACA 1399
Db 661 TGAATCCAGATATAGGGAGAAAAATGAAAGATATCTATTGTGTCTTTTCAATGATGAACA 720
Qy 1400 CCATTCATATGACCACTCATATACAGCTTCAAAAGAGCTTTGACTGTGAGCTCGCAGA 1459
Db 721 CCATTCATATGACCACTCATATACAGCTTCAAAAGAGCTTTGACTGTGAGCTCGCAGA 780
Qy 1460 GGCCAGTGTGATACCACTGCATTCGATCAAAAGAGGTCGTGGGCTGTAAAGCGGGAGC 1519
Db 781 GGCCAGTGTGATACCACTGCATTCGATCAAAAGAGGTCGTGGGCTGTAAAGCGGGAGC 840
Qy 1520 TTATGTCTCTTGGCCAGGAGCAAGAGGAGATATAAAGAGTCAATTCAGAAAAATGTCTCTCA 1579
Db 841 TTATGTCTCTTGGCCAGGAGCAAGAGGAGATATAAAGAGTCAATTCAGAAAAATGTCTCTCA 900
Qy 1580 ACATCCACTTCATGTAGAAAGTATTACACTCAGAGATTAATGGCTCATCAGAAATTTGCTTT 1639
Db 901 ACATCCACTTCATGTAGAAAGTATTACACTCAGAGATTAATGGCTCATCAGAAATTTGCTTT 960
Qy 1640 GCGTCTTGGTTCCTGGATGAACAAAAATATGAGCTATTCAAGTGAATTTAGGCGAGATCTT 1699
Db 961 GCGTCTTGGTTCCTGGATGAACAAAAATATGAGCTATTCAAGTGAATTTAGGCGAGATCTT 1020
Qy 1700 TTGCCAAGCATGCCCTTAGAGAAGAACCTGACTCGGAGAAATCCCTGTCTCATAGCAGGTT 1759
Db 1021 TTGCCAAGCATGCCCTTAGAGAAGAACCTGACTCGGAGAAATCCCTGTCTCATAGCAGGTT 1080
Qy 1760 AATGCTTTGGATGCAAAAGCTTTATAAAGGTGCCGTAAGATCCCTTCATGAATTTGATCTT 1819
Db 1081 AATGCTTTGGATGCAAAAGCTTTATAAAGGTGCCGTAAGATCCCTTCATGAATTTGATCTT 1140
Qy 1820 CAGCAGTTTTTTTTTATGGAGATGGAATATACAAAAACTCTTTCGTATGGAATTTGGAAGTA 1879
Db 1141 CAGCAGTTTTTTTTTATGGAGATGGAATATACAAAAACTCTTTCGTATGGAATTTGGAAGTA 1200
Qy 1880 TTATAACAACTCGCAAGAAAGATATATCAGTGTATGATCATGACAGAGTATCTCTATAAC 1939
Db 1201 TTATAACAACTCGCAAGAAAGATATATCAGTGTATGATCATGACAGAGTATCTCTATAAC 1260
Qy 1940 TGCATTTTCAGTTCAGATGTTTTACTGTTCTCTATCTCGGCTCGACATCTTATGAGAGCA 1999
Db 1261 TGCATTTTCAGTTCAGATGTTTTACTGTTCTCTACTCTCGGCTCGACATCTTATGAGAGCA 1320

Qy 2000 GAATGTTATCTCTGTCAATTACTGAAACTCTGCTAGAGTTTTTACCTGAGTACTTGGACAG 2059
Db 1321 GAATGTTATCTCTGTCAATTACTGAAACTCTGCTAGAGTTTTTACCTGAGTACTTGGACAG 1380
Qy 2060 GAACAATAAATTCAACTTCAGGGTTATAGCCAGGACAAATTTGGGAAAGATATATGCAGT 2119
Db 1381 GAACAATAAATTCAACTTCAGGGTTATAGCCAGGACAAATTTGGGAAAGATATATGCAGT 1440
Qy 2120 AATATGTGACCTAAAGTATATCTGATCAGCAAAACCCACAAATATGGACAGAAATTAAG 2179
Db 1441 AATATGTGACCTAAAGTATATCTGATCAGCAAAACCCACAAATATGGACAGAAATTAAG 1500
Qy 2180 AATGTCAGTTCCTTGAAGGTTTTCGATCTTTTTTTGAAGATCTTTTACCTGTATGCAAGGAAT 2239
Db 1501 AATGTCAGTTCCTTGAAGGTTTTCGATCTTTTTTTGAAGATCTTTTACCTGTATGCAAGGAAT 1560
Qy 2240 GGAAGAAATCCGAAGACAGGTTGGGCAACACATTTGAAGTGGATCCTGATTTGGGAGGCTGC 2299
Db 1561 GGAAGAAATCCGAAGACAGGTTGGGCAACACATTTGAAGTGGATCCTGATTTGGGAGGCTGC 1620
Qy 2300 CATTTGCTATACAGATGCAATTTGAAGATATTTTACTCATGTTCCAAAGAGTGGTGTGCTTG 2359
Db 1621 CATTTGCTATACAGATGCAATTTGAAGATATTTTACTCATGTTCCAAAGAGTGGTGTGCTTG 1680
Qy 2360 TGATGAAGAACTCTTACTTGTGGCTTATAAAGAAATGTCAAAAGCTGTGATGAGGTGCAG 2419
Db 1681 TGATGAAGAACTCTTACTTGTGGCTTATAAAGAAATGTCAAAAGCTGTGATGAGGTGCAG 1740
Qy 2420 TACAGTTTCATATCTATAGTAGCAAGACAGTAGTATCAATCTGTTGGTGGACATAGTTGGGAAAC 2479
Db 1741 TACCAGTTTCATATCTATAGTAGCAAGACAGTAGTATCAATCTGTTGGTGGACATAGTTGGGAAAC 1800
Qy 2480 AAAGTCTTACAGAGTATCTGAGGATCTTGTAAAGCATACATCTGCCATCTCTAGAGCCCT 2539
Db 1801 AAAGTCTTACAGAGTATCTGAGGATCTTGTAAAGCATACATCTGCCATCTCTAGAGCCCT 1860
Qy 2540 TGCTGGTCTTTCATGTACGTTTTAAGCAGGCTGGGGTGTGTTTTCAAGACTGCAATGAATTTGT 2599
Db 1861 TGCTGGTCTTTCATGTACGTTTTAAGCAGGCTGGGGTGTGTTTTCAAGACTGCAATGAATTTGT 1920
Qy 2600 GTCTTTTGGAGCACTTTCAGAGTAGAGGTACTAGTGGAAATATCCCTTTTACGTTGCTGGTGT 2659
Db 1921 GTCTTTTGGAGCACTTTCAGAGTAGAGGTACTAGTGGAAATATCCCTTTTACGTTGCTGGTGT 1980
Qy 2660 GGTGGCCAGGTTGTTGCTCAGATGTGGCGAAGAAATGGACTGTCTCTTATATGACAGGTT 2719
Db 1981 GGTGGCCAGGTTGTTGCTCAGATGTGGCGAAGAAATGGACTGTCTCTTATATGACAGGTT 2040
Qy 2720 GTTTTATTACCAAGATGTTAAGTGCAGAGAAATGTATGATATAAGATATCATCATGCT 2779
Db 2041 GTTTTATTACCAAGATGTTAAGTGCAGAGAAATGTATGATATAAGATATCATCATGCT 2100
Qy 2780 TCAGATTTGGTGCATCTTTAATTTGAATGCCAATAAAGTCTTGTGTACTTGGTACTTCAGAGGTA 2839
Db 2101 TCAGATTTGGTGCATCTTTAATTTGAATGCCAATAAAGTCTTGTGTACTTGGTACTTCAGAGGTA 2160
Qy 2840 TGAATTTGCCGAGGCTTTTAAACAGACCATATCTACAAAGACCAAGGATTTGATTAACA 2899
Db 2161 TGAATTTGCCGAGGCTTTTAAACAGACCATATCTACAAAGACCAAGGATTTGATTAACA 2220
Qy 2900 ATATAATACACTTAATAGAAGAAATGCTTTCAGGTCTCATCTATATTTGGGTGAGCGTTA 2959
Db 2221 ATATAATACACTTAATAGAAGAAATGCTTTCAGGTCTCATCTATATTTGGGTGAGCGTTA 2280
Qy 2960 TGTAACCTGGAGTGGGAAATGTGACCAAGAGAGGTCACAAATGAGAGAAATCATTCACCT 3019
Db 2281 TGTAACCTGGAGTGGGAAATGTGACCAAGAGAGGTCACAAATGAGAGAAATCATTCACCT 2340
Qy 3020 GCTTTTGCATTTGAACCCATGCCACAGTGGCCATTTGCCAAAAATTTTACTCTGAGAGATGAAA 3079
Db 2341 GCTTTTGCATTTGAACCCATGCCACAGTGGCCATTTGCCAAAAATTTTACTCTGAGAGATGAAA 2400

QY 3080 TAATGAAATGGCTTAGAGATGTCTATAAACAAGTGGCCACATTTTAAAGAACCAAGGTGT 3139

Db 2401 TAATGAAATGGCTTAGAGATGTCTATAAACAAGTGGCCACATTTTAAAGAACCAAGGTGT 2460

QY 3140 ATCAGGCCATGGAGTTTATGAATCAATGAAGTGAATCACTGAAGACTTCAATATATGACTT 3199

Db 2461 ATCAGGCCATGGAGTTTATGAATCAATGAAGTGAATCACTGAAGACTTCAATATGACTT 2520

QY 3200 TTATCATTTCTCAAAACCCAGCATAGCAAGGCTGAACATATGCGAGAGAAAAGAGAAA 3259

Db 2521 TTATCATTTCTCAAAACCCAGCATAGCAAGGCTGAACATATGCGAGAGAAAAGAGAAA 2580

QY 3260 ACAAGAAACAAGATGAAGCATTCGCCGCCACACCTCTCTGAATTTCTGCCCTGCTTT 3319

Db 2581 ACAAGAAACAAGATGAAGCATTCGCCGCCACACCTCTCTGAATTTCTGCCCTGCTTT 2640

QY 3320 CAGCAAGTGAATTAACCTTTCTCAACTGTGATATCATGATGATATCTCAGGACCGTATT 3379

Db 2641 CAGCAAGTGAATTAACCTTTCTCAACTGTGATATCATGATGATATCTCAGGACCGTATT 2700

QY 3380 TGAGCGGCATAGACACAGATTTCTAATCTGTGGACCGAAGGATGCTCCAAATGGCTTT 3439

Db 2701 TGAGCGGCATAGACACAGATTTCTAATCTGTGGACCGAAGGATGCTCCAAATGGCTTT 2760

QY 3440 TCATATTTCTGGCATTTGGCTTTACTAGAGAGAGCAACAGCTTCAAAAGCTCCTGAAGA 3499

Db 2761 TCATATTTCTGGCATTTGGCTTTACTAGAGAGAGCAACAGCTTCAAAAGCTCCTGAAGA 2820

QY 3500 AGAAGTAACTTTGATTTTATCATAGGCTTCAAGATTTGGGAAGTCTCAGGCATGAATAT 3559

Db 2821 AGAAGTAACTTTGATTTTATCATAGGCTTCAAGATTTGGGAAGTCTCAGGCATGAATAT 2880

QY 3560 ACAATGCTTTTGGAAAACCTCAAGGAATTTCCCGAGTTAGAGGCCAGAGGACATGAT 3619

Db 2881 ACAATGCTTTTGGAAAACCTCAAGGAATTTCCCGAGTTAGAGGCCAGAGGACATGAT 2940

QY 3620 AAGCTGGATCTTCAGATGTTTTCACACAGTGAAGCGATTAGAGAAAATCTTGTGTTAAT 3679

Db 2941 AAGCTGGATCTTCAGATGTTTTCACACAGTGAAGCGATTAGAGAAAATCTTGTGTTAAT 3000

QY 3680 TGTAGCAACCATCATCAGGATCGGAATCTATTAAAGATGATGATGATGATGATGATGATGAT 3738

Db 3001 TGTAGCAACCATCATCAGGATCGGAATCTATTAAAGATGATGATGATGATGATGATGATGAT 3059

RESULT 11

LOCUS AK056441

DEFINITION Homo sapiens cDNA FLJ31879 fis, clone NT2RP7002619, highly similar to Mus musculus ubiquitin-protein ligase E3-alpha (Ubr1) mRNA.

ACCESSION AK056441

VERSION AK056441.1 GI:16551844

KEYWORDS oligo capping; fis (full insert sequence).

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Wakatsuki, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H., Sekine, M., Ohtsuka, M., Nishi, T., Shibahara, T., Tanaka, T., Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y., Nagahara, K., Murakami, K., Tanai, H., Iwayanagi, T., Wagatsuma, M., Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T., Kikawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H., Hiraoka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S., Yosida, M., Hotuka, T., Kusanagi, J., Kanehori, K., Takahashi-Fujii, A., Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R., Takeuchi, K., Arita, M., Imose, N., Mueshino, K., Yuuki, H., Oshima, A., Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T.,

Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S., Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H., Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T., Sugiyama, A., Takemoto, B., Kawakami, B., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y., Ozaki, K., Hirao, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Ohtani, R., Kawakami, T., Noguchi, S., Iton, T., Shigeta, K., Senba, T., Matsumura, K., Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togaishi, T., Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J., Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K., Negase, T., Nomura, N., Kikuchi, H., Maehuo, Y., Yamashita, R., Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S. Complete sequencing and characterization of 21,243 full-length human cDNAs Nat. Genet. 36 (1), 40-45 (2004)

2 Tashiro, H., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakatsuki, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahara, K., Maehuo, Y., Nagai, K. and Isogai, T. NEDO human cDNA sequencing project Unpublished

3 (bases 1 to 3059)

Isogai, T., Otsuki, T. and Sugiyama, T. Direct Submission

Submitted (24-Oct-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@hri.co.jp. Tel: 81-438-52-3975, Fax: 81-438-52-3986)

NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: RAB and HRI.

FEATURES

source

1. 3059

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="NT2RP7002619"

/cell_line="NT2"

/clone_lib="NT2RP7"

/note="cloning vector: pME18SPL3-mRNA from NT2 neuronal precursor cells after 5-weeks retinoic acid (RA) induction."

ORIGIN

Query Match 48.4%; Score 3055.8; DB 9; Length 3059;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 3057; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 680 GTCAGTTTCCCTCAAGATGGCGGACGAGGAGGCTGGAGGTACTGAGAGATGGAATCAG 739

Db 1 GTCAGTTTCCCTCAAGATGGCGGACGAGGAGGCTGGAGGTACTGAGGAGATGGAATCAG 60

QY 740 CGCGAGTACCCAGACCCCTTCAGCGCTCTGGCATCTTGGGGATCAGCAAGTTGATT 799

Db 61 CGCGAGTACCCAGACCCCTTCAGCGCTCTGGCATCTTGGGGATCAGCAAGTTGATT 120

QY 800 TTATCTGCTTCTTGTGATCATTTTGGACAAATTTGGTGCAGAAATTTTACTTTGCTGAAT 859

Db 121 TTATCTGCTTCTTGTGATCATTTTGGACAAATTTGGTGCAGAAATTTTACTTTGCTGAAT 180

QY 860 GGACCCAGACTTGGAAAGCAGGAGGAAGTGTACAAATGTCAATATTTCACTCCACTGGA 919
DB 181 GGAACCCAGACTTGGAAAGCAGGAGGAAGTGTACAAATGTCAATATTTCACTCCACTGGA 240
QY 920 ATGCTACTTATTTGGAGAGAGTCCAGATATTTGCTTAGAGAAATTTGAAGCAGTGGAGC 979
DB 241 ATGCTACTTATTTGGAGAGAGTCCAGATATTTGCTTAGAGAAATTTGAAGCAGTGGAGC 300
QY 980 ATTTTCAGCTTTGCTGGAGGGTTTTCAAAAGTGGAGAGACACCTATTTCTTGGAGGATTTG 1039
DB 301 ATTTTCAGCTTTGCTGGAGGGTTTTCAAAAGTGGAGAGACACCTATTTCTTGGAGGATTTG 360
QY 1040 TGCATATGATCCAACTGTGACTCTGTATGGAAGTCTGCTTCCAGGACAGTGTTCATAAAAA 1099
DB 361 TGCATATGATCCAACTGTGACTCTGTATGGAAGTCTGCTTCCAGGACAGTGTTCATAAAAA 420
QY 1100 TCATCGTTTACAAGATGCACTTCTTACTTGGAGGAGGGTTCTGTGACTGTGGAGACACAGA 1159
DB 421 TCATCGTTTACAAGATGCACTTCTTACTTGGAGGAGGGTTCTGTGACTGTGGAGACACAGA 480
QY 1160 GGATGGAAACTGGCCCTTTTGTGTAAATCATGAACCTGGAGAGAGAGTACTATAAA 1219
DB 481 GGATGGAAACTGGCCCTTTTGTGTAAATCATGAACCTGGAGAGAGTACTATAAA 540
QY 1220 AGAGAAATTCAGCTGTCCGTTGAAATGAAGAGGTAAATGTCGAAGCCAGGAAATATTTCC 1279
DB 541 AGAGAAATTCAGCTGTCCGTTGAAATGAAGAGGTAAATGTCGAAGCCAGGAAATATTTCC 600
QY 1280 TTCAGTGATAAAATATGTCGTAGAAATGATATATGGGAAGGAGGAAAGAACTGCTTCC 1339
DB 601 TTCAGTGATAAAATATGTCGTAGAAATGATATATGGGAAGGAGGAAAGAACTGCTTCC 660
QY 1340 TGAATCCAGATAGGAGGAAATGAAGAGTACTTATGTCCTTTTCAATGATGAACA 1399
DB 661 TGAATCCAGATAGGAGGAAATGAAGAGTACTTATGTCCTTTTCAATGATGAACA 720
QY 1400 CCATTTCAATACACAGCTCATATACAGCTTCAAAAGAGCTCTGACTGTGAGCTCGAGA 1459
DB 721 CCATTTCAATACACAGCTCATATACAGCTTCAAAAGAGCTCTGACTGTGAGCTCGAGA 780
QY 1460 GGCCCGAGTTGCATACCACTGCCATTTGCAAAAGAGGGTCTCGGGCTGTTAAAGCGGAGC 1519
DB 781 GGCCCGAGTTGCATACCACTGCCATTTGCAAAAGAGGGTCTCGGGCTGTTAAAGCGGAGC 840
QY 1520 TTATGCTGCTTCCGAGGAGCAAGAGAGATATAAAGAGTCAATTCAGAAATGTCTCA 1579
DB 841 TTATGCTGCTTCCGAGGAGCAAGAGAGATATAAAGAGTCAATTCAGAAATGTCTCA 900
QY 1580 ACATCCACTTCATGTAGAAGTATTACACTCAGAGATTTATGCTCATCAGAAATTTGCTTT 1639
DB 901 ACATCCACTTCATGTAGAAGTATTACACTCAGAGATTTATGCTCATCAGAAATTTGCTTT 960
QY 1640 CGCTTCTGGTTCTCGATGAACAAATTTATGAGCTATTTCAAGTGAATTTAGGAGATCTT 1699
DB 961 CGCTTCTGGTTCTCGATGAACAAATTTATGAGCTATTTCAAGTGAATTTAGGAGATCTT 1020
QY 1700 TTGCCAAGCTGCTTAGAGAGAACCTGACTCGGAGATCCCTGTCTCATAGCAGGTT 1759
DB 1021 TTGCCAAGCTGCTTAGAGAGAACCTGACTCGGAGATCCCTGTCTCATAGCAGGTT 1080
QY 1760 AATGCTTTGGAGTGCAGAGCTTTATAAGGTGCGGTAAAGTCTTTCATGAATTTGATCTT 1819
DB 1081 AATGCTTTGGAGTGCAGAGCTTTATAAGGTGCGGTAAAGTCTTTCATGAATTTGATCTT 1140
QY 1820 CAGCAGTTTTTTTATGGAGATGGATACAAAAAATCTTTGCTATGGAAATTTGGAAGTA 1879
DB 1141 CAGCAGTTTTTTTATGGAGATGGATACAAAAAATCTTTGCTATGGAAATTTGGAAGTA 1200
QY 1880 TTATAACCACTGCAGAAAGATATATCAGTGATGATCATCAGAGAGTATCTCTATAAC 1939
DB 1201 TTATAACCACTGCAGAAAGATATATCAGTGATGATCATCAGAGAGTATCTCTATAAC 1260
QY 1940 TGCACCTTTTCAGTTCAGATGTTTACTGTTCTTACTCTGCTCGACATCTTATTTGAAGAGCA 1999

DB 1261 TGCACCTTTTCAGTTCAGATGTTTACTGTTCTCTACTCTGCTCGACATCTTATTTGAAGAGCA 1320
QY 2000 GAATGTTATCTCTGTCTCAATCTGAACTCTGCTAGAGTCTTACCTCAGTACTTTGGACAG 2059
DB 1321 GAATGTTATCTCTGTCTCAATCTGAACTCTGCTAGAGTCTTACCTCAGTACTTTGGACAG 1380
QY 2060 GAACAATAAATTTCAACTTTCCAGGGTTATAGCCAGGACAAATTTGGGAAGAGTATATGCACT 2119
DB 1381 GAACAATAAATTTCAACTTTCCAGGGTTATAGCCAGGACAAATTTGGGAAGAGTATATGCACT 1440
QY 2120 AATATGTACCTAAAGATATATCTGTATCAGCAAAACCACAAATATGGACAGAAAGTAAAG 2179
DB 1441 AATATGTACCTAAAGATATATCTGTATCAGCAAAACCACAAATATGGACAGAAAGTAAAG 1500
QY 2180 AATGCACTTCTTGAAGGTTTTCGATCTTTTGAAGATCTTACCTGTATGCAAGGCAAT 2239
DB 1501 AATGCACTTCTTGAAGGTTTTCGATCTTTTGAAGATCTTACCTGTATGCAAGGCAAT 1560
QY 2240 GGAAGAAATCCGAAGACAGGTTGGGCAACATTTGAAGTGGATCTCTGATTTGGGAGGCTGC 2299
DB 1561 GGAAGAAATCCGAAGACAGGTTGGGCAACATTTGAAGTGGATCTCTGATTTGGGAGGCTGC 1620
QY 2300 CATTTGCTATACAGATGCAATTTGAAGATATTTTACTCATGTTCACAGAGTGGTGTCTTG 2359
DB 1621 CATTTGCTATACAGATGCAATTTGAAGATATTTTACTCATGTTCACAGAGTGGTGTCTTG 1680
QY 2360 TGATGAAGAACTCTTACTTGTGCTTATAAGATGTCANAGCTGTGATGAGGCTGCAG 2419
DB 1681 TGATGAAGAACTCTTACTTGTGCTTATAAGATGTCANAGCTGTGATGAGGCTGCAG 1740
QY 2420 TACCAGTTTTCATATCTAGTAGCAAGACAGTAGTAGTACAATCGTGTGCAATAGTTTGGAAAC 2479
DB 1741 TACCAGTTTTCATATCTAGTAGCAAGACAGTAGTAGTACAATCGTGTGCAATAGTTTGGAAAC 1800
QY 2480 AAAGTCTCTACAGAGTACTCAGAGATCTTGAAGATACATCTGCACTCTCTAGGACCT 2539
DB 1801 AAAGTCTCTACAGAGTACTCAGAGATCTTGAAGATACATCTGCACTCTCTAGGACCT 1860
QY 2540 TGCTGCTTTCATGCTAGCTTTAAGCAGGCTGGGTGCTGTTTCAAGATGCAATGATTTGT 2599
DB 1861 TGCTGCTTTCATGCTAGCTTTAAGCAGGCTGGGTGCTGTTTCAAGATGCAATGATTTGT 1920
QY 2600 GTCTTTTGGAGACTTTTCAAGTAGAGTACTAGTGGAAATATCTTTTACTGCTGTGTT 2659
DB 1921 GTCTTTTGGAGACTTTTCAAGTAGAGTACTAGTGGAAATATCTTTTACTGCTGTGTT 1980
QY 2660 GGTTGCCCGAGGTTGCTCTGAGATGTCGCGAAGAAATGGACTGTCTCTTATTTAGCCAGGT 2719
DB 1981 GGTTGCCCGAGGTTGCTCTGAGATGTCGCGAAGAAATGGACTGTCTCTTATTTAGCCAGGT 2040
QY 2720 GTTTTATTTACAGATGTTAAGTGCAGAGAGAAATGTTATGATTAAGATATCATCATGCT 2779
DB 2041 GTTTTATTTACAGATGTTAAGTGCAGAGAGAAATGTTATGATTAAGATATCATCATGCT 2100
QY 2780 TCAGATTTGGTGCATCTTTTAAATGGAATCCAAATAGTCTTTTACTGCTTCTCAGAGGTA 2839
DB 2101 TCAGATTTGGTGCATCTTTTAAATGGAATCCAAATAGTCTTTTACTGCTTCTCAGAGGTA 2160
QY 2840 TGAACCTTCCCGAGGCTTTTAAACAGACCATATCTCAAAAAGACCCAGGATTTGATTAACA 2899
DB 2161 TGAACCTTCCCGAGGCTTTTAAACAGACCATATCTCAAAAAGACCCAGGATTTGATTAACA 2220
QY 2900 ATATAATACACTAATAGAGAAATGCTTCAAGTCTCTCATCTATTTTGGGTGAGCGTTA 2959
DB 2221 ATATAATACACTAATAGAGAAATGCTTCAAGTCTCTCATCTATTTTGGGTGAGCGTTA 2280
QY 2960 TGTACCTGGAGTGGAAATGTGACAAAGAGAGGTCAATATGAGAGAAATCATTCATT 3019
DB 2281 TGTACCTGGAGTGGAAATGTGACAAAGAGAGGTCAATATGAGAGAAATCATTCATT 2340
QY 3020 GCTTTTGCATTTGAACCCCATGCGCACAGTGGCAATTTGCGCAAAATTTTACTGAGATGAAAA 3079

Db 2341 GCTTTGCTTGAACCCATGCCACACAGTGCCTATTCGCCAAAATTTTACCTGAGATGAAGA 2400
Qy 3080 TAATGAATCTGCTTTAGAGATGTCTATAAACAAGTGGCCACATTTAAGAAACCGAGTGT 3139
Db 2401 TAATGAATCTGCTTTAGAGATGTCTATAAACAAGTGGCCACATTTAAGAAACCGAGTGT 2460
Qy 3140 ATCAGGCCATGAGTTTATGAACTAAAGATGAATCACTGAAGACTTCAATATGTACTTT 3199
Db 2461 ATCAGGCCATGAGTTTATGAACTAAAGATGAATCACTGAAGACTTCAATATGTACTTT 2520
Qy 3200 TTATCATTTACTCCAAAACCCAGCATAGCAAGGCTGAACATATGCAAGAAAGAGAGAAA 3259
Db 2521 TTATCATTTACTCCAAAACCCAGCATAGCAAGGCTGAACATATGCAAGAAAGAGAGAAA 2580
Qy 3260 ACAAGAAACAAAGATGAAGCATTCGCCGCCACACACCTCTGTAATTTCTGCCCTGCTTT 3319
Db 2581 ACAAGAAACAAAGATGAAGCATTCGCCGCCACACACCTCTGTAATTTCTGCCCTGCTTT 2640
Qy 3320 CAGCAAGTGAATTAACCTTCTCACTGATATCATGATGTACATTTCTCAGGACCGTATT 3379
Db 2641 CAGCAAGTGAATTAACCTTCTCACTGATATCATGATGTACATTTCTCAGGACCGTATT 2700
Qy 3380 TGAGCGGCAATAGACACAGATTTCTAATCTTGTGACCGAAGGGATGCTCCAAATGGCTTT 3439
Db 2701 TGAGCGGCAATAGACACAGATTTCTAATCTTGTGACCGAAGGGATGCTCCAAATGGCTTT 2760
Qy 3440 TCATATTTCTGCGATTTGCTTTTACAGAGAGAGCAACAGCTTCAAAAAGCTTCTGGAAGA 3499
Db 2761 TCATATTTCTGCGATTTGCTTTTACAGAGAGAGCAACAGCTTCAAAAAGCTTCTGGAAGA 2820
Qy 3500 AGAAGTACATTTGATTTTATCATAGGCTTCAAGATTTGGAAGTTTCAAGCATGAATAT 3559
Db 2821 AGAAGTACATTTGATTTTATCATAGGCTTCAAGATTTGGAAGTTTCAAGCATGAATAT 2880
Qy 3560 ACAATGCTTTTGGAAAACCTCAAGGAATTTCCCGAGTTAGAGGCCAAGAGGACATCAT 3619
Db 2881 ACAATGCTTTTGGAAAACCTCAAGGAATTTCCCGAGTTAGAGGCCAAGAGGACATCAT 2940
Qy 3620 AACGTGGATCTCAGATGTTTGACACAGTGAAGCGATTAAGAGAAAATCTTGTGTTAAT 3679
Db 2941 AACGTGGATCTCAGATGTTTGACACAGTGAAGCGATTAAGAGAAAATCTTGTGTTAAT 3000
Qy 3680 TGTAGCACCATCAGATCGGATCTATTAAAGATCATGATGATTAATCATGATAAAG 3738
Db 3001 TGTAGCACCATCAGATCGGATCTATTAAAGATCATGATGATTAATCATGATAAAG 3059

RESULT 12

BD156870
LOCUS BD156870 2550 bp DNA linear PAT 17-JAN-2003
DEFINITION Primer for synthesizing full-length cDNA and use thereof.

ACCESSION BD156870

VERSION BD156870.1 GI:27862628

KEYWORDS JP 2002191363-A/11713.

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 2550)

Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J.,

Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.

Primer for synthesizing full-length cDNA and use thereof

Patent: JP 2002191363-A 11713 09-JUL-2002;

HELIX RESEARCH INSTITUTE

OS Homo sapiens (human)

PN JP 2002191363-A/11713

PD 09-JUL-2002

PF 28-JUL-2000 JP 2000280990

PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU

PI SAITO,

PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,

PI KEIICHI NAGAI,TETSUJI OTSUKI

PC

C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/PC
10,
PC C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC

Primer for synthesizing full-length cDNA and use thereof FH Key
Location/Qualifiers
(117) (2549).

FT CDS

Location/Qualifiers

source

1..2550

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match 40.4%; Score 2546.8; DB 6; Length 2550;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 2548; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1189 ATCATGAACCTGGAGAGCAGTACTATAAAGAGAAATTCACGCTGTCGCTTGAATGAAG 1248

Db 1 ATCATGAACCTGGAGAGCAGTACTATAAAGAGAAATTCACGCTGTCGCTTGAATGAAG 60

Qy 1249 AGGTAATTTGTCACAGCCAGGAAATATTTCTTTCAGTGAATAAATATGCTAGAGAAATGA 1308

Db 61 AGGTAATTTGTCACAGCCAGGAAATATTTCTTTCAGTGAATAAATATGCTAGAGAAATGA 120

Qy 1309 CTATATGGAAGAGAGAAAGAACTGCTCTCTGAACTCCAGATAGAGGAGAGAAATGAAA 1368

Db 121 CTATATGGAAGAGAGAAAGAACTGCTCTCTGAACTCCAGATAGAGGAGAGAAATGAAA 180

Qy 1369 GATACTATTTGTCCTTTTCAATGATGAACCACTTCAATATGACCACTCATATACAGCC 1428

Db 181 GATACTATTTGTCCTTTTCAATGATGAACCACTTCAATATGACCACTCATATACAGCC 240

Qy 1429 TACAAGAGCTCTTGAAGTGTGAGCTCGCAGAGGCCAGTTCATATGACCACTGCAATTTGACA 1488

Db 241 TACAAGAGCTCTTGAAGTGTGAGCTCGCAGAGGCCAGTTCATATGACCACTGCAATTTGACA 300

Qy 1489 AAGAGGTGCTGCGGCTGTTAAAGCGGAGCTTATGCTTTCAGAGGAGGAGGAGGAGGAG 1548

Db 301 AAGAGGTGCTGCGGCTGTTAAAGCGGAGCTTATGCTTTCAGAGGAGGAGGAGGAGGAG 360

Qy 1549 ATATAAGAGTCATTCAGAAATGCTCTCAACATCCACTTCAATGATGAGGATTTACACT 1608

Db 361 ATATAAGAGTCATTCAGAAATGCTCTCAACATCCACTTCAATGATGAGGATTTACACT 420

Qy 1609 CAGAGATTTAGCTCATCAGAAATTTGCTTTGCGTCTTGGTTCCTGGATGAACAAATTA 1668

Db 421 CAGAGATTTAGCTCATCAGAAATTTGCTTTGCGTCTTGGTTCCTGGATGAACAAATTA 480

Qy 1669 TGAGCTATTCAAGTGACTTTAGGAGATCTTTTCCAGCATGCTTTAGAGAGAACCTG 1728

Db 481 TGAGCTATTCAAGTGACTTTAGGAGATCTTTTCCAGCATGCTTTAGAGAGAACCTG 540

Qy 1729 ACTCGGAGATCCCTGCTCATAGAGAGGTTAATGCTTTGGGATGCAAGCTTTATAAG 1788

Db 541 ACTCGGAGATCCCTGCTCATAGAGAGGTTAATGCTTTGGGATGCAAGCTTTATAAG 600

Qy 1789 GTGCCCGTAAGATCCTTCATGAATTTGATCTTTCAGCAGTCTTTTATGAGATGGAATACA 1848

Db 601 GTGCCCGTAAGATCCTTCATGAATTTGATCTTTCAGCAGTCTTTTATGAGATGGAATACA 660

Qy 1849 AAAAAGCTTTTGTCTATGGAATTTGTAAGTATTATAAACAACCTGAGAGAAAGATATATCA 1908

Db 661 AAAAAGCTTTTGTCTATGGAATTTGTAAGTATTATAAACAACCTGAGAGAAAGATATATCA 720

Qy 1909 GTGATGATCATGACAGAGATATCTCTATACTGCACTTTTCAGTTCAGATGTTTACTGTTTC 1968

Db 721 GTGATGATCATGACAGAGATATCTCTATACTGCACTTTTCAGTTCAGATGTTTACTGTTTC 780

Qy 1969 CTACTCTGGCTCGACATCTTATTAAGAGCAGAGATTTATCTCTGCTCAATTAAGAACTC 2028

Db 781 CTACTCTGGCTCGACATCTTATTAAGAGCAGAGATTTATCTCTGCTCAATTAAGAACTC 840

Qy	2029	TGCTAGAGT	TTTTTACCTGAGT	ACTTGGACAGGAA	CAATAAATTTCAACTTCCAGGGTTATA	2089
Db	841	TGCTAGAGT	TTTTTACCTGAGT	ACTTGGACAGGAA	CAATAAATTTCAACTTCCAGGGTTATA	900
Qy	2089	GCCAGGACA	AAATTTGGGAAGAGT	ATATGCCAGT	TAATATGTGACCTTAAGTATATCTCTGATCA	2148
Db	901	GCCAGGACA	AAATTTGGGAAGAGT	ATATGCCAGT	TAATATGTGACCTTAAGTATATCTCTGATCA	960
Qy	2149	GCAAACTCCACA	ATATGACAGAAAGATTAAAGAT	TGCAGTTTCCCTTGAAGGT	TTTTTTCGATCTT	2208
Db	961	GCAAACTCCACA	ATATGACAGAAAGATTAAAGAT	TGCAGTTTCCCTTGAAGGT	TTTTTTCGATCTT	1020
Qy	2209	TTTTTGAAGATTCTT	TACCTGTATGACGGGAAT	TGGAAGAAATCCGAAGACAGGTTGGGCAAC	2268	
Db	1021	TTTTTGAAGATTCTT	TACCTGTATGACGGGAAT	TGGAAGAAATCCGAAGACAGGTTGGGCAAC	1080	
Qy	2269	ACATTGAAGTGGAT	TCCTGTATTTGGGAGGCTGCCAT	TGCTATACAGATGCAATTTGAAGAATA	2328	
Db	1081	ACATTGAAGTGGAT	TCCTGTATTTGGGAGGCTGCCAT	TGCTATACAGATGCAATTTGAAGAATA	1140	
Qy	2329	TTTTTACTCATGTTTCCAAGAGTGGT	TGCTTGTGATGAAGAACTTCTTACTTTGCGCTTTATA	2388		
Db	1141	TTTTTACTCATGTTTCCAAGAGTGGT	TGCTTGTGATGAAGAACTTCTTACTTTGCGCTTTATA	1200		
Qy	2389	AAGAAATGTC	CAAAAGCTGTGATGAGGTG	GCAGTACCAAGTTTTCATATCTAGTAGCAAGACAG	2448	
Db	1201	AAGAAATGTC	CAAAAGCTGTGATGAGGTG	GCAGTACCAAGTTTTCATATCTAGTAGCAAGACAG	1260	
Qy	2449	TAGTACAATCGTGTGGACA	TAGTTTGGAAA	CAAAGTCCATACAGAGTATCTGAGGATCTTGG	2508	
Db	1261	TAGTACAATCGTGTGGACA	TAGTTTGGAAA	CAAAGTCCATACAGAGTATCTGAGGATCTTGG	1320	
Qy	2509	TAAGCATACATCGCCACTCT	TAGGACCCCTGCTGTCCTTCAATGACGTTTAAAGCAGGC	2568		
Db	1321	TAAGCATACATCGCCACTCT	TAGGACCCCTGCTGTCCTTCAATGACGTTTAAAGCAGGC	1380		
Qy	2569	TGGGTGCTGTTTCAAAGACTGCATGAA	TTTGTGCTTTTGGAGGACTTTTCAAGTAGAGGTAC	2628		
Db	1381	TGGGTGCTGTTTCAAAGACTGCATGAA	TTTGTGCTTTTGGAGGACTTTTCAAGTAGAGGTAC	1440		
Qy	2629	TAGTGGAAATATCCTTTACGTTTGTGTTGTGGT	TGCCAAGTTGTTGTCGAGATGTGGC	2688		
Db	1441	TAGTGGAAATATCCTTTACGTTTGTGTTGTGGT	TGCCAAGTTGTTGTCGAGATGTGGC	1500		
Qy	2689	GAAGAAATGCACTGCTCTTATTTAGCCAGGTGTTTATTAACCAAGATGTTAAAGTCAGAG	2748			
Db	1501	GAAGAAATGCACTGCTCTTATTTAGCCAGGTGTTTATTAACCAAGATGTTAAAGTCAGAG	1560			
Qy	2749	AAGAAATGTTATGATAAAGATATCATCATGCTTCAGATTGCTGTCATCTTTAATGGATCCCA	2808			
Db	1561	AAGAAATGTTATGATAAAGATATCATCATGCTTCAGATTGCTGTCATCTTTAATGGATCCCA	1620			
Qy	2809	ATAAGTTCTTGTTTACTGGTACTTCAGAGGTTATGAACTTTCGCGAGGCTTTTAAACAAGACCA	2868			
Db	1621	ATAAGTTCTTGTTTACTGGTACTTCAGAGGTTATGAACTTTCGCGAGGCTTTTAAACAAGACCA	1680			
Qy	2869	TATCTCAAAAGACCGAGTTTGATTTAAACAATATATACACTAATAGAGAAATGCTTC	2928			
Db	1681	TATCTCAAAAGACCGAGTTTGATTTAAACAATATATACACTAATAGAGAAATGCTTC	1740			
Qy	2929	AGTTCCTCATCTATTTTGTGGGTGAGCGTTATGTACCTCGAGTGGGAAATGTGACCAAAG	2988			
Db	1741	AGTTCCTCATCTATTTTGTGGGTGAGCGTTATGTACCTCGAGTGGGAAATGTGACCAAAG	1800			
Qy	2989	AAGAGGTACAAATGAGAGAAATCATCTTGTCTTGCATTTGAACCCATGCCACAAGTG	3048			
Db	1801	AAGAGGTACAAATGAGAGAAATCATCTTGTCTTGCATTTGAACCCATGCCACAAGTG	1860			
Qy	3049	CCATTGCCAAAAATTTTACCTGAGAAATGAATAATGAACTGCGCTTAGAGAAATGCTATAA	3108			
Db	1861	CCATTGCCAAAAATTTTACCTGAGAAATGAATAATGAACTGCGCTTAGAGAAATGCTATAA	1920			
Qy	3109	ACAAAGTGGCCACATTTTAAAGAAACCAAGTGTATCAGGCCATGGAGTTTATGCAACTAAAG	3168			

Db	1921	ACAAAGTGGCCACATTTAAAGAAACCAAGGTGTATCAGGCCATGGAGTTTATGAACATAAAG	1980
Qy	3169	ATGAATCAGCTGAAGAGCTTCAATATGTACTTTTATCATTAATCTCCAAACCCAGCATAGCA	3228
Db	1981	ATGAATCAGCTGAAGAGCTTCAATATGTACTTTTATCATTAATCTCCAAACCCAGCATAGCA	2040
Qy	3229	AGGCTGAACATATGCAAGAAAGGAGAAACAAAGAAACAAAGATGAAGCAATGCGCG	3288
Db	2041	AGGCTGAACATATGCAAGAAAGGAGAAACAAAGAAACAAAGATGAAGCAATGCGCG	2100
Qy	3289	CACCAACACCTCTGAATTCCTGCGCTGCTTTTCAGCAAAAGTGATTAACCTTCTCAACTGTG	3348
Db	2101	CACCAACACCTCTGAATTCCTGCGCTGCTTTTCAGCAAAAGTGATTAACCTTCTCAACTGTG	2160
Qy	3349	ATATCATGATGTACATTTCTCAGGACCGTATTTTCAGCGGGCAATAGACACAGATTTCTAACT	3408
Db	2161	ATATCATGATGTACATTTCTCAGGACCGTATTTTCAGCGGGCAATAGACACAGATTTCTAACT	2220
Qy	3409	TGTGGACCGAAGGGATGCTCCAAATGGCTTTTCATATTCCTGGCATTTGGGTTTACTAGAAG	3468
Db	2221	TGTGGACCGAAGGGATGCTCCAAATGGCTTTTCATATTCCTGGCATTTGGGTTTACTAGAAG	2280
Qy	3469	AGAAGCAACAGCTTCAAAAAGCTCTCGAAGAGAAGTAACATTTTGACTTTTATCATTAAG	3528
Db	2281	AGAAGCAACAGCTTCAAAAAGCTCTCGAAGAGAAGTAACATTTTGACTTTTATCATTAAG	2340
Qy	3529	CTTTCAGATTTGGGAAGTTTCAGGCCATGAATATACAAATGCTTTTGGAAAACTCAAAGGAA	3588
Db	2341	CTTTCAGATTTGGGAAGTTTCAGGCCATGAATATACAAATGCTTTTGGAAAACTCAAAGGAA	2400
Qy	3589	TTCCCCAGTTAGAAGGCCAGAGGACATGATAACCTGATGATTTTCAGATGTTTGACACAG	3648
Db	2401	TTCCCCAGTTAGAAGGCCAGAGGACATGATAACCTGATGATTTTCAGATGTTTGACACAG	2460
Qy	3649	TGAAGCGATTAGAAGAAAAATCTTGTGTTAATTTGTAGCAACACATCAGGATCGGAATCTA	3708
Db	2461	TGAAGCGATTAGAAGAAAAATCTTGTGTTAATTTGTAGCAACACATCAGGATCGGAATCTA	2520
Qy	3709	TTAAGATGATGAGATTACTCATGATAAAG	3738
Db	2521	TTAAGATGATGAGATTACTCATGATAAAG	2550
RESULT	13		
AX877826			
LOCUS	AX877826	2550 bp	DNA
DEFINITION	Sequence 12731 from Patent EP1074617.		linear
ACCESSION	AX877826		
VERSION	AX877826.1	GI:40032562	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE			
AUTHORS	Oka, T., Isogai, T., Nishikawa, T., Hayashi, K., Saico, K., Yamamoto, J., Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Oseuki, T.		
TITLE	Primers for synthesising full-length cDNA and their use		
JOURNAL	Patent: EP 1074617-A 12731 07-FEB-2001.		
FEATURES	Research Association for Biotechnology (JP)		
source	Location/Qualifiers		
	1..2550		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
	117..>2549		
	/note="unnamed protein product"		
	/codon_start=1		
	/protein_id="CA892293.1"		
	/db_xref="GI:40032563"		
	/translation="MTIWEBEKSLPPELQIREKNRYCYVLFNDEHSHYDHVITYSLROR		
	ALPCEIABAEQLHTTADIKGERRAVTKAGAYAACAKEDIKSHSNVSOHPLHVEVLHS		
CDS			

EIMAHOKPALRLGSMWNKIMSYSSDFPROIFCOACILRBPDSNPCLISRLMLWDKLY
KARLILHELIFSPMEBYKGFAMBEFYKIQOLQKXEYISDDHRSISITALSVMQ
FTVPLALHELIEBQNVISVITETLLEVLPEYLDNRNKNFQYSDQKGRVYAVICDL
KYILSKPTIWTWRMQLBGRFLKILTCMOGMBEIRQVGHIEVDWDEAIIA
IOMOLKNTLLAFQECACDEBELLVAYKECHKAVNRCSFTSSSTVVQSGHLE
KSYRVEDLVS IHLPLSRTLAGHLVRLSRLGAVSRLEHBFVDFEDFOVEVLVEYPLRCL
VYVAQVAMWRRENGLSLSISQVFFYQDVKREEMVDKDIIMLOIGASLMDPNKELLV
LQRYELAPFKNTISTKODPLIKQNTLIEMLQVLIYIVGERIYVPGVGNVKEBTV
REIHLHLCIEPMPHSAIAKNUPENNETGLEINKVATPKPGVSGHYVELKDES
LKDPMFYHYHSKQHSKAEHMKRKRQENKDEALPPPPPEFPAFSGVYNLLNCD
IMWYILTRVFERAIDTDSNLWTEGMLQMAFHILALGLLEEKQOLKAKABEVEVDFYH
KASRLGSSAMNIOMLEKLGIPOLEGQKDMITWILQMPDFTVRLREKSLIVTSG
SESINKDEITHDK"

ORIGIN

Query Match 40.4%; Score 2546.8; DB 6; Length 2550;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 2548; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	1189	ATCATGAACCTGGAAGAGCAGTACTATATAAGAGAAATTCACGCTGTCGCTTGAATGAAG	1248
DB	1	ATCATGAACCTGGAAGAGCAGTACTATATAAGAGAAATTCACGCTGTCGCTTGAATGAAG	60
QY	1249	AGGTAATTTGCCAAGCCAGGAAATATTTCTTCAGTGTATAAATATGTCCTAGAAATGA	1308
DB	61	AGGTAATTTGCCAAGCCAGGAAATATTTCTTCAGTGTATAAATATGTCCTAGAAATGA	120
QY	1309	CTATATGGAAGAGGAGAAAGAACTGCGCTCCTGAACTCAGATAAGGAGAGAAATGAAA	1368
DB	121	CTATATGGAAGAGGAGAAAGAACTGCGCTCCTGAACTCAGATAAGGAGAGAAATGAAA	180
QY	1369	GATACATTTGTGCTCTTTCAATGANGAACACCATTCATATGACCACTCATATACAGCC	1428
DB	181	GATACATTTGTGCTCTTTCAATGANGAACACCATTCATATGACCACTCATATACAGCC	240
QY	1429	TACAAGAGCTCTTGACCTGAGCTCGCAGAGGCCAGTTGTCATACCACTGCCATTGACA	1488
DB	241	TACAAGAGCTCTTGACCTGAGCTCGCAGAGGCCAGTTGTCATACCACTGCCATTGACA	300
QY	1489	AAGAGGCTCGTCGGGCTGTTAAGCGGGAGCTTATGCTGCTTGCAGGAAGCAAGAGGAG	1548
DB	301	AAGAGGCTCGTCGGGCTGTTAAGCGGGAGCTTATGCTGCTTGCAGGAAGCAAGAGGAG	360
QY	1549	ATATAAGAGTCATTACAGAAATGCTCTCAACATCCACTTCATGATAGAAATTTACACT	1608
DB	361	ATATAAGAGTCATTACAGAAATGCTCTCAACATCCACTTCATGATAGAAATTTACACT	420
QY	1609	CAGAGATTTAGCTCATCAGAAATTTGCTTTGGCTTTGGTCTTGGATGAACAAATTA	1668
DB	421	CAGAGATTTAGCTCATCAGAAATTTGCTTTGGCTTTGGTCTTGGATGAACAAATTA	480
QY	1669	TGAGCTATTCAAGTGACTTTAGGCGAGATCTTTTGCAGAGCATGCTTACAGAGAACCTG	1728
DB	481	TGAGCTATTCAAGTGACTTTAGGCGAGATCTTTTGCAGAGCATGCTTACAGAGAACCTG	540
QY	1729	ACTCGAGATCCCTGTCTCATAGCAGGTTAATGCTTTGGGATGCAAGCTTTATAAAG	1788
DB	541	ACTCGAGATCCCTGTCTCATAGCAGGTTAATGCTTTGGGATGCAAGCTTTATAAAG	600
QY	1789	GTGCCGCTAGATCTTCATGAATTTGATCTTCAGCAGGTTTTTTTATGAGATGGATACA	1848
DB	601	GTGCCGCTAGATCTTCATGAATTTGATCTTCAGCAGGTTTTTTTATGAGATGGATACA	660
QY	1849	AAAAACTCTTGTCTATGAAATTTGGAAGTATTATAAACAACCTGACAGAAAGATATATCA	1908
DB	661	AAAAACTCTTGTCTATGAAATTTGGAAGTATTATAAACAACCTGACAGAAAGATATATCA	720
QY	1909	GTGATGATCATGACAGAGATCTCTATAACTGCACCTTTTCAGTTTCAGATGTTTACTGTT	1968
DB	721	GTGATGATCATGACAGAGATCTCTATAACTGCACCTTTTCAGTTTCAGATGTTTACTGTT	780
QY	1969	CTACTCTGGCTCGACATCTTATTTGAAGAGCAGAAATGTTTATCTGTCTATTACTGAACTC	2028

DB	781	CTACTCTGGCTCGACATCTTATTGAAGAGCAGAAATGTTTATCTCTGTCTATTACTGAACTC	840
QY	2029	TGCTAGAAAGTTTTACCTGAGTACTTTGGACAGGAAACAATAAATTTCAATTTCCAGGGTTATA	2088
DB	841	TGCTAGAAAGTTTTACCTGAGTACTTTGGACAGGAAACAATAAATTTCAATTTCCAGGGTTATA	900
QY	2089	GCCAGGACAAATTTGGGAAGAGTATATGACGTAATAATGTGACCTAAAGTATATCTCTGATCA	2148
DB	901	GCCAGGACAAATTTGGGAAGAGTATATGACGTAATAATGTGACCTAAAGTATATCTCTGATCA	960
QY	2149	GCAAAACCAACAATATGGAACAGAAATTAAGAAATGACGTTCTTCTGAAAGGTTTTCCATCTT	2208
DB	961	GCAAAACCAACAATATGGAACAGAAATTAAGAAATGACGTTCTTCTGAAAGGTTTTCCATCTT	1020
QY	2209	TTTTGAAGATTTCTTACCTGTATGACGGAATGGAAGAAATCCGAAGACAGGTTGGGCAAC	2268
DB	1021	TTTTGAAGATTTCTTACCTGTATGACGGAATGGAAGAAATCCGAAGACAGGTTGGGCAAC	1080
QY	2269	ACATTTGAAGTGGATCCTGNTTTGGGAGGCTGCCATTTGCTATACAGATGCAATTTGAAGATA	2328
DB	1081	ACATTTGAAGTGGATCCTGNTTTGGGAGGCTGCCATTTGCTATACAGATGCAATTTGAAGATA	1140
QY	2329	TTTTTACTCATCTTCCAGAGTGTGTCTTGTGATGAAGAACTCTTACTTTGGGCTTATA	2388
DB	1141	TTTTTACTCATCTTCCAGAGTGTGTCTTGTGATGAAGAACTCTTACTTTGGGCTTATA	1200
QY	2389	AAGAAATGTCAAAAGCTGTGTATGAGGTGACGATCCAGTTTCTATCTAGTAGCAGACAG	2448
DB	1201	AAGAAATGTCAAAAGCTGTGTATGAGGTGACGATCCAGTTTCTATCTAGTAGCAGACAG	1260
QY	2449	TAGTACAATCTGTGTGACATATGTTTGGAAACAAAGTCTCTACAGATATCTCAGAGATCTTG	2508
DB	1261	TAGTACAATCTGTGTGACATATGTTTGGAAACAAAGTCTCTACAGATATCTCAGAGATCTTG	1320
QY	2509	TAAGCATACATCTGCCACTCTCTAGGACCTTCTAGGACCTTCTGCTTCTCATGATGAGGAGC	2568
DB	1321	TAAGCATACATCTGCCACTCTCTAGGACCTTCTAGGACCTTCTGCTTCTCATGATGAGGAGC	1380
QY	2569	TGGGTGCTGTTTCAAGACTGATGAAATTTGTGTCTTTTGGAGACTTTTCAAGTAGAGGTAC	2628
DB	1381	TGGGTGCTGTTTCAAGACTGATGAAATTTGTGTCTTTTGGAGACTTTTCAAGTAGAGGTAC	1440
QY	2629	TAGTGGAAATATCTTTTACGTTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	2688
DB	1441	TAGTGGAAATATCTTTTACGTTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1500
QY	2689	GAAGAAATGAGTCTCTCTTATTAGCCAGGTTTATTATTAACAAGATGTTAAGTGCACAG	2748
DB	1501	GAAGAAATGAGTCTCTCTTATTAGCCAGGTTTATTATTAACAAGATGTTAAGTGCACAG	1560
QY	2749	AAGAAATGTATGATAAAGATATCATCTGCTTACAGATTTGGTGCATCTTTAATGATCCCA	2808
DB	1561	AAGAAATGTATGATAAAGATATCATCTGCTTACAGATTTGGTGCATCTTTAATGATCCCA	1620
QY	2809	ATAAGTTCTTGTTACTGGTACTTCAGAGGTATGAACCTTTCGCGAGGCTTTTAAACAGACA	2868
DB	1621	ATAAGTTCTTGTTACTGGTACTTCAGAGGTATGAACCTTTCGCGAGGCTTTTAAACAGACA	1680
QY	2869	TATCTCAAAAGACAGGATTTGATTAACAATAATATACACTAATAAGAAATGCTTTC	2928
DB	1681	TATCTCAAAAGACAGGATTTGATTAACAATAATATACACTAATAAGAAATGCTTTC	1740
QY	2929	AGGTCTCTCATCTATTTTGTGGGTGAGGCTTATGTATGACCTGAGGTGGAAATGTGACCAAG	2988
DB	1741	AGGTCTCTCATCTATTTTGTGGGTGAGGCTTATGTATGACCTGAGGTGGAAATGTGACCAAG	1800
QY	2989	AAGAGGTCACAATGAGAGAAATCATCTCACTTGTCTTGTGCAATTTGAACCCCATGCAAGTG	3048
DB	1801	AAGAGGTCACAATGAGAGAAATCATCTCACTTGTCTTGTGCAATTTGAACCCCATGCAAGTG	1860
QY	3049	CCATTGCCAAAAATTTTACCTGAGATGAAATTAATAATGAACTGGCTTAGAGAAATGTCTATA	3108
DB	1861	CCATTGCCAAAAATTTTACCTGAGATGAAATTAATAATGAACTGGCTTAGAGAAATGTCTATA	1920

QY 3109 ACAAGTGGCCACATTTTAAGAAACACAGGTGTATCAGGCCATGGAGTGTATTAAGTAACTAAAG 3168
Db 1921 ACAAGTGGCCACATTTTAAGAAACACAGGTGTATCAGGCCATGGAGTGTATTAAGTAACTAAAG 1980
QY 3169 ATGAATCACTGAAGACTTCAATATGTACTTTTATCATTACTCCAAAACCCAGCATAGCA 3228
Db 1981 ATGAATCACTGAAGACTTCAATATGTACTTTTATCATTACTCCAAAACCCAGCATAGCA 2040
QY 3229 AGGCTGAATATGACAGAAAGAGAGAAACAGAAACAAAGATGAAGCATGGCCGC 3288
Db 2041 AGGCTGAATATGACAGAAAGAGAGAAACAAAGATGAAGCATGGCCGC 2100
QY 3289 CACCAACCACTCTCAATTTCTGCCCTGCTTTCAGCAAAAGTGAATTAACCTTCTCAACTGTG 3348
Db 2101 CACCAACCACTCTCAATTTCTGCCCTGCTTTCAGCAAAAGTGAATTAACCTTCTCAACTGTG 2160
QY 3349 ATATCATGATGTACATTTCTCAGGACCGGTATTTGAGCGGGCAATAGACACAGATTTCAACT 3408
Db 2161 ATATCATGATGTACATTTCTCAGGACCGGTATTTGAGCGGGCAATAGACACAGATTTCAACT 2220
QY 3409 TGTGACCGAAGGATGCTCCAAATGGCTTTTCATATTTCTGCGCATGTTGGTTCATAGAG 3468
Db 2221 TGTGACCGAAGGATGCTCCAAATGGCTTTTCATATTTCTGCGCATGTTGGTTCATAGAG 2280
QY 3469 AGAAGCAACAGCTTCAAAAGCTCTGAGAGAGATTAACATTTGACATTTTATCATAGG 3528
Db 2281 AGAAGCAACAGCTTCAAAAGCTCTGAGAGAGATTAACATTTGACATTTTATCATAGG 2340
QY 3529 CTTCAAGATTGGGAAGTTTCAGCCATGAATATACAAATGCTTTTGAAAGAACTCAAAGGA 3588
Db 2341 CTTCAAGATTGGGAAGTTTCAGCCATGAATATACAAATGCTTTTGAAAGAACTCAAAGGA 2400
QY 3589 TTCCCAAGTTAGAGCCAGAGGACATGATTAACCTGATATCTTCAGATGTTTGACACAG 3648
Db 2401 TTCCCAAGTTAGAGCCAGAGGACATGATTAACCTGATATCTTCAGATGTTTGACACAG 2460
QY 3649 TGAAGCGATTAGAGAAAATCTGTTTAATGTTAGCAACCAATCAGATCGGAATCTA 3708
Db 2461 TGAAGCGATTAGAGAAAATCTGTTTAATGTTAGCAACCAATCAGATCGGAATCTA 2520
QY 3709 TTAAGATGATGAGATTAATCTCATGATAAG 3738
Db 2521 TTAAGATGATGAGATTAATCTCATGATAAG 2550
RESULT 14
AK027803
LOCUS
DEFINITION Homo sapiens cDNA FLJ14897 fig. clone PLACE1004743, weakly similar to PROBABLE N-END-RECOGNIZING PROTEIN.
ACCESSION AK027803
VERSION AK027803.1 GI:14042751
KEYWORDS oligo capping; fig (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1
AUTHORS Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H., Sekine, M., Ohyashi, M., Nishi, T., Shibahara, T., Tanaka, T., Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y., Nagahara, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M., Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T., Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuma, N., Sato, K., Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H., Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M., Hirooka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S., Yosida, M., Hotta, T., Kaneshiro, K., Takahashi, F., Fujii, A., Hara, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R., Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuuki, H., Oshima, A.,

Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T., Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S., Teraehima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H., Goto, Y., Shimizu, P., Wakebe, H., Hishigaki, H., Watanabe, T., Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y., Ozaki, K., Hirao, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T., Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsumura, K., Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T., Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J., Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K., Nagase, T., Nomura, K., Kikuchi, H., Masuho, Y., Yamashita, R., Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S.
Complete sequencing and characterization of 21,243 full-length human cDNAs
Nat. Genet. 36 (1), 40-45 (2004)

2
Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y., Nishikawa, T., Nagai, K., Sugano, S., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagahara, K., Masuho, Y. and Sasaki, N.
NEDO human cDNA sequencing project
Unpublished
3 (bases 1 to 2550)
Isogai, T. and Otsuki, T.
Direct Submission
Submitted (10-MAY-2001) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction: 5' - 3' - end one pass sequencing and clone selection: Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.

FEATURES
source

1. .2550
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE1004743"
/tissue_type="placenta"
/clone_lib="PLACE1"
/note="cloning vector: pME18SFL3"
117. .>2550
/note="unnamed protein product"
/codon_start=1
/protein_id="BAB55380.1"
/db_xref="GI:14042752"
/translation="MTIWEKEKLPPELQIRKERNRYCYLVFNDEHSHYDHYVLSQRL
ALDCELAQHLHTTAIDKRRRAVAGAYAAACQAKEDIKSHSENVSQHPHLHVLHLS
EIMAHQKFAIRLGLSMWNKIMSYSDSPQIFCOACLRBEPSENCPLISRLMLDKLVL
KGAPKILHELIFSSFFMEYKCLFAMEFVKYKQLOKEYISDDHRSISITALSVMQ
FTVPTLARIHEQNIVSVITETILLEVLPEVLDNRNKNFQYSGDQLGVVYICDL
KYILISKPTIWTIRMQFLGFRSFKLILTCMGMBEIRKQVQGHIEVDWRAAIA
IQMKNLMLFQWCACDELLVAYKCHKAVMRCSTSPISSTKTVQSCGHSLET
LQVQVAEMWRNGLSLISQVFFYQDVCKREMYDKDIIMLGASIMDNPKEVLV
LVRYELAEAFNKTTSTKQVNTLIREMLQVLIYIGERVGVGVVYKKEVTVM
RELIIHLICBEPHSAIAKXLPENNETGLENVKAVFKPGVSGHGVYELKDES
LKDPMYFHYKSTQHSKARQKRNQENKORALPPPPPPCFQAPKAPKSVINLLNC
IMVILTVFERAIDTDSNLWTEGMLQMAFHILALGLLEKEQQLCPQAKBEBVTDFYH
KASLGGSSAMNIQMLEKLGIPQLEGQKDMITWILQMFDTVKELREKSLIVATTSG
SESINKDEIITHDK"

CDS

ORIGIN

Query Match 40.4%; Score 2546.8; DB 9; Length 2550;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2548; Conservative 0; Mismatches 2; Indels 0; Gaps 0;


```
|||||
2161 ATATCATGATGTACATTTCTCAGGACCGTATTTGAGCGGCAATAGACACAGATTTCTAACT 2220
QY
3409 TGTGACCGAAGGATGCTCCAAATGGCTTTTCATATTTCTGGCATGGGTTTACTAGAAG 3468
Db
2221 TGTGACCGAAGGATGCTCCAAATGGCTTTTCATATTTCTGGCATGGGTTTACTAGAAG 2280
QY
3469 AGAAGCAACAGCTTCAAAAGCTCTGGAAGAGTAACATTTTGACATTTTATCATAAAG 3528
Db
2281 AGAAGCAACAGCTTCAAAAGCTCTGGAAGAGTAACATTTTGACATTTTATCATAAAG 2340
QY
3529 CTTCAAGATTTGGAAGTTTCAGCCATGAATATACAAATGCTTTTGGAAAAAATCTCAAAAGAA 3588
Db
2341 CTTCAAGATTTGGAAGTTTCAGCCATGAATATACAAATGCTTTTGGAAAAAATCTCAAAAGAA 2400
QY
3589 TTCCCCAGTTAGAAGCCGAGAAGGACATGATTAACCTGGATATCTTCAGATGTTTGAACAG 3648
Db
2401 TTCCCCAGTTAGAAGCCGAGAAGGACATGATTAACCTGGATATCTTCAGATGTTTGAACAG 2460
QY
3649 TGAAGCGATTGAAGAAAAATCTGTTTAAATTTGTAGCAACACATCAGGATCGGAATCTA 3708
Db
2461 TGAAGCGATTGAAGAAAAATCTGTTTAAATTTGTAGCAACACATCAGGATCGGAATCTA 2520
QY
3709 TTAAGAATGATGAGATTAATCTCATGATAAAG 3738
Db
2521 TTAAGAATGATGAGATTAATCTCATGATAAAG 2550

RESULT 15
HSM808548 4576 bp mRNA linear PRI 30-AUG-2003
LOCUS Homo sapiens mRNA; cDNA DKFZp686M07224 (from clone DKFZp686M07224).
DEFINITION BX648400
ACCESSION BX648400.1 GI:34367562
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4576)
AUTHORS Bloesker, H., Boecker, M., Mewes, H.W., Weil, B., Amid, C., Osanger, A.,
Fobo, G., Han, M. and Wiemann, S.
CONSTRM The German Human cDNA Consortium
TITLE Direct Submission
JOURNAL Submitted (27-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by GBF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKFZp686M07224) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.
FEATURES
Location/Qualifiers
source
1..4576
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686M07224"
/tissue_type="human small intestine"
/clone_lib="686 (synonym: hicc3). Vector pSport1_Sfi; host
DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
polya_signal 4528..4533
polya_site 4549
ORIGIN
Query Match 39.4%; Score 2487.8; DB 9; Length 4576;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2489; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY
3818 GAAAAAATTTCATTTGAAAACTCATAACTCATGTATGACAAATACATACAGAAATGCCTGGAA 3877
Db
1 GAAAAAATTTCATTTGAAAACTCATAACTCATGTATGACAAATACATACAGAAATGCCTGGAA 60
QY
3878 AGAAGATTCCATTTATGAGGAAGAGAGACACCCAGCAGTCACTGACACTCTAGAAATGC 3937
Db
61 AGAAGATTCCATTTATGAGGAAGAGAGAGACACCCAGCAGTCACTGACACTCTAGAAATGC 120
QY
3938 TTTGGGTCTTAAACGGGGTCCATCTGTTACTGAAAAAGAGGTGCTGAGCTGCATCCTTTG 3997
Db
121 TTTGGGTCTTAAACGGGGTCCATCTGTTACTGAAAAAGAGGTGCTGAGCTGCATCCTTTG 180
QY
3998 CCAAGAAAGAACAGCAGAGTGAATAATAGAAAAATAATGCCATGTATTTATCGGCTGTGTCCA 4057
Db
181 CCAAGAAAGAACAGCAGAGTGAATAATAGAAAAATAATGCCATGTATTTATCGGCTGTGTCCA 240
QY
4058 GAAATCTACTGCTTTAAACCCAGACACAGGGGAAAAACCCATAGAACTCTCAGGAGAAGCCCT 4117
Db
241 GAAATCTACTGCTTTAAACCCAGACACAGGGGAAAAACCCATAGAACTCTCAGGAGAAGCCCT 300
QY
4118 AGACCCACTTTTCATGATCCAGACTTTGGCATATGGAACCTTATACAGGAAGCTGTGTCA 4177
Db
301 AGACCCACTTTTCATGATCCAGACTTTGGCATATGGAACCTTATACAGGAAGCTGTGTCA 360
QY
4178 TGTAAATGCACGAGTGTGCTGGCAGAAAGTATTTTGAAGCTGTACAGCTGAGCTCTCAGCA 4237
Db
361 TGTAAATGCACGAGTGTGCTGGCAGAAAGTATTTTGAAGCTGTACAGCTGAGCTCTCAGCA 420
QY
4238 GCGCAATTCATGTTGACCTTTTGAAGTGGGAGAAATATCTTTGCCCTCTTTGCAA 4297
Db
421 GCGCAATTCATGTTGACCTTTTGAAGTGGGAGAAATATCTTTGCCCTCTTTGCAA 480
QY
4298 ATCTCTGTGCAATATCTGTGATCCCATATTTCTTTGCAACCTCAAAAGATAAACAGTGA 4357
Db
481 ATCTCTGTGCAATATCTGTAAATCTTCTTTGCAACCTCAAAAGATAAACAGTGA 540
QY
4358 GAATGCAGATGCTCTGCTCAACTTTTGAACCTGGCAGGTGGATACAGACTGTTTGGC 4417
Db
541 GAATGCAGATGCTCTGCTCAACTTTTGAACCTGGCAGGTGGATACAGACTGTTTGGC 600
QY
4418 CAGAAATTCAGGTTTAAATATAAGACATGCTTAAAGGAGAAAAACCAATTCCTATTTCTT 4477
Db
601 CAGAAATTCAGGTTTAAATATAAGACATGCTTAAAGGAGAAAAACCAATTCCTATTTCTT 660
QY
4478 TAATCAAGGAATGGGAGATTTCTACTTTGGAGTTCCATTCCTCTGAGTTTGGGCTTGA 4537
Db
661 TAATCAAGGAATGGGAGATTTCTACTTTGGAGTTCCATTCCTCTGAGTTTGGGCTTGA 720
QY
4538 GTCTTCGATTAAATATTTCAAAATAGCATCAAGGAATGTTATTTCTTTTGGCAACAACAT 4597
Db
721 GTCTTCGATTAAATATTTCAAAATAGCATCAAGGAATGTTATTTCTTTTGGCAACAACAT 780
QY
4598 TTATAGAAATTTGGAATGGAAGTCCACCTGATGAAAGGGATCCTCGAGTCCCATCTGAC 4657
Db
781 TTATAGAAATTTGGAATGGAAGTCCACCTGATGAAAGGGATCCTCGAGTCCCATCTGAC 840
QY
4658 CTGAGACACCTGCGCTTTCTACTATCCAGGCAATTTGAAAAATCTATTTGGAGATGAAGGAA 4717
Db
841 CTGAGACACCTGCGCTTTCTACTATCCAGGCAATTTGAAAAATCTATTTGGAGATGAAGGAA 900
QY
4718 ACCTCTGTTTGGAGCACTTTCAAAATAGGCATTAATGTTCTGAAAGCATTTAATGCAAGTT 4777
Db
901 ACCTCTGTTTGGAGCACTTTCAAAATAGGCATTAATGTTCTGAAAGCATTTAATGCAAGTT 960
QY
4778 TGCAGTTGCAACAGAGGATTAATCTGCTCTCAGGTCTCTGATACAGAAACATCTGGTTCGCT 4837
Db
961 TGCAGTTGCAACAGAGGATTAATCTGCTCTCAGGTCTCTGATACAGAAACATCTGGTTCGCT 1020
QY
4838 TCTATCAGTTGTTCTTCTTAACATAAAATCAGAAAGATACACCATGCTTCTGCTATAGA 4897
Db
1021 TCTATCAGTTGTTCTTCTTAACATAAAATCAGAAAGATACACCATGCTTCTGCTATAGA 1080
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 8, 2005, 18:33:52 ; Search time 2893 Seconds
(without alignments)
12907.614 Million cell updates/sec

Title: US-10-758-672A-1
Perfect score: 6308
Sequence: 1 gccagagattggcagcaggg.....aattttgtattgtgtgtttt 6308

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2359870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6308	100.0	6308	12 ADJ95441	Adj95441 Human cDN
2	6308	100.0	6308	13 ADS86863	ADS86863 cDNA enco
3	5592	88.6	7742	8 ABZ24689	ABZ24689 Human cel
4	5388	85.4	7648	12 ADI16377	ADI16377 Human pro
5	5136	81.4	5205	12 ADJ95458	ADJ95458 Human cDN
6	5136	81.4	5205	13 ADS86880	ADS86880 cDNA enco
7	4435.4	70.3	6395	2 AAV99308	AAV99308 cDNA enco
8	4435.4	70.3	6395	5 AAC86933	AAC86933 Nucleotid
9	3055.8	48.4	3059	10 ADA53348	ADA53348 Human cod
10	2546.8	40.4	2550	4 AAH14878	AAH14878 Human cDN
11	1716.4	27.2	3776	13 ACN39798	ACN39798 Tumour-as
12	1529.2	24.2	3598	13 ADR07317	ADR07317 Full leng
13	1381.8	21.9	6300	12 ADJ95443	ADJ95443 Human cDN
14	1381.8	21.9	6300	13 ADS86865	ADS86865 cDNA enco
15	1311	20.8	1635	6 ABQ75898	ABQ75898 Human ubi
16	1233.6	19.6	6089	12 ADJ95445	ADJ95445 Mouse cDN
17	1233.6	19.6	6089	13 ADS86867	ADS86867 cDNA enco
18	1099.4	17.4	6840	4 AAK51709	AAK51709 Human pol
19	1090	17.3	6850	4 AAK52693	AAK52693 Human pol
20	996.2	15.8	1001	2 AAX03300	AAX03300 Partial c

21	996.2	15.8	1001	5 AAC86934	AAC86934 Nucleotid
22	981	15.6	6158	13 ADR14633	ADR14633 Human NF-
23	813.6	12.9	3327	2 AAX35730	AAX35730 cDNA enco
24	813.6	12.9	3502	2 AAX35731	AAX35731 cDNA enco
25	797.2	12.6	4208	12 ADQ64363	ADQ64363 Novel hum
26	656.8	10.4	733	3 AAA02411	AAA02411 Human col
27	647	10.3	712	4 AAH07621	AAH07621 Human cDN
28	638.6	10.1	756	3 AAA02327	AAA02327 Human col
29	626.8	9.9	2616	6 ABZ70161	ABZ70161 Fibrinoge
30	450.8	7.1	777	5 AAS70068	AAS70068 DNA enco
31	438.4	6.9	505	9 ACH49509	ACH49509 Human leu
32	429	6.8	1464	5 AAS70104	AAS70104 DNA enco
33	374.8	5.9	1772	6 ABL90577	ABL90577 Human pol
34	294.4	4.7	4073	13 ADR07162	ADR07162 Full leng
35	271.4	4.3	1172	5 AAS70851	AAS70851 DNA enco
36	250.6	4.0	907	6 ABK35394	ABK35394 Human cDN
37	246.8	3.9	802	2 AAX39891	AAX39891 Gastric c
38	241.6	3.8	264	3 AAA00233	AAA00233 Human col
39	219.8	3.5	4573	8 ACA03886	ACA03886 cDNA upre
40	216.8	3.4	6278	4 ABL14799	ABL14799 Drosophil
c 41	191.6	3.0	800	2 AAX39892	AAX39892 Gastric c
c 42	160.2	2.5	3411	10 ADB62029	ADB62029 Human cDN
c 43	159	2.5	455	2 AAV89674	AAV89674 EST clone
c 44	142	2.3	643	6 ABL89546	ABL89546 Human pol
c 45	136.6	2.2	550	12 ACH79702	ACH79702 Human gen

ALIGNMENTS

RESULT 1

ADJ95441
ID ADJ95441 standard; cDNA; 6308 BP.

XX AC ADJ95441;

DT 03-JUN-2004 (first entry)

XX DE Human cDNA encoding Ubiquitin ligase E3alpha I.

XX KW Human; ss; gene; ubiquitin ligase; E3alpha I;
KW ubiquitin-proteasome pathway; gene therapy; vaccine; muscular atrophy;
KW cachexia; catabolic disorders; cancer cachexia; renal cachexia;
KW inflammatory cachexia; muscle wasting disorder; metabolic acidosis;
KW uremia; burn; hyperthyroidism; Cushing's syndrome; fasting;
KW denervation atrophy; diabetes mellitus; sepsis; AIDS wasting syndrome;
KW SNP; single nucleotide polymorphism.

XX OS Homo sapiens.

XX FH Key
XX CDS Location/Qualifiers
696..5945

XX FT /*tag= a
XX FT /product= "Ubiquitin ligase E3alpha I"
XX FT /replace= (5397,C)
XX FT /*tag= b
XX FT /standard_name= "Single nucleotide polymorphism"

XX US6706505-B1.

XX PD 16-MAR-2004.

XX PF 28-NOV-2000; 2000US-00724126.

XX PR 08-MAR-2000; 2000US-0187911P.

XX PA (AMGE-) AMGEN INC.

XX PI Han H, Kwak K;

XX DR WPI; 2004-236723/22.

XX DR P-PSDB; ADJ95442.

XX

1741	DB	CCTGTCCTCATGAAGCAGGTTAATGCTTTGGGANGCAAGCTTTATAAAGGTGGCCGGTAAGA	1800
1801	QY	TCCTTCATGAATTCGATCTTCAGCAGTGTTCCTTTATGAGAGATGGAATACAAAAAATCTTTTG	1860
1801	DB	TCCTTCATGAATTCGATCTTCAGCAGTGTTCCTTTATGAGAGATGGAATACAAAAAATCTTTTG	1860
1861	QY	CTATGGAAATTGTGAAGTATTTATTAACAACCTGCAGAAAGAAATATATCAGTGTGATGATCATG	1920
1861	DB	CTATGGAAATTGTGAAGTATTTATTAACAACCTGCAGAAAGAAATATATCAGTGTGATGATCATG	1920
1921	QY	ACAGAAGTATCTCTATAACTGCACCTTCCTCAGTTCAGATGTTTACTGTTCTCTACTCTGGCTC	1980
1921	DB	ACAGAAGTATCTCTATAACTGCACCTTCCTCAGTTCAGATGTTTACTGTTCTCTACTCTGGCTC	1980
1981	QY	GACATCTTATGGAAGACAGAAATGTTATCTCTGTCAATTACTGAAATCTCTCTAGAAGTTT	2040
1981	DB	GACATCTTATGGAAGACAGAAATGTTATCTCTGTCAATTACTGAAATCTCTCTAGAAGTTT	2040
2041	QY	TACCTGAGTACTTGGACAGGAAACAATAAATTCAACTTCAGGGTTATAGCCAGGACAAAT	2100
2041	DB	TACCTGAGTACTTGGACAGGAAACAATAAATTCAACTTCAGGGTTATAGCCAGGACAAAT	2100
2101	QY	TGGGAAGAGTATATGCAAGTATATGTCACCTTAAGTATATCCTGATCAGCAAAACCCACAA	2160
2101	DB	TGGGAAGAGTATATGCAAGTATATGTCACCTTAAGTATATCCTGATCAGCAAAACCCACAA	2160
2161	QY	TATGGACAGAAAGATTAAAGATTCAGTTCCTTGAAGGTTTTCGATCTTTTTCGAAGATTC	2220
2161	DB	TATGGACAGAAAGATTAAAGATTCAGTTCCTTGAAGGTTTTCGATCTTTTTCGAAGATTC	2220
2221	QY	TTACTCTGATGCAAGGAAATGGAAGAAATCCGAAGACAGGTTGGGCAACAATTTGAAGTGG	2280
2221	DB	TTACTCTGATGCAAGGAAATGGAAGAAATCCGAAGACAGGTTGGGCAACAATTTGAAGTGG	2280
2281	QY	ATCCTGATTCGGAGGCTGCCATTGCTATACAGATGCAATTTGAAGATATTTTACTCATGT	2340
2281	DB	ATCCTGATTCGGAGGCTGCCATTGCTATACAGATGCAATTTGAAGATATTTTACTCATGT	2340
2341	QY	TCCAAGAGTGGTGTGCTGTGTGATGAAGAACTCTTACTTGTGGCTTATAAAGAATGTCAACA	2400
2341	DB	TCCAAGAGTGGTGTGCTGTGTGATGAAGAACTCTTACTTGTGGCTTATAAAGAATGTCAACA	2400
2401	QY	AAGCTGTGATGAGGTGCAATCCAGTTTCATATCTAGTAGCAAGACAGTAGTAGTACAATCGT	2460
2401	DB	AAGCTGTGATGAGGTGCAATCCAGTTTCATATCTAGTAGCAAGACAGTAGTAGTACAATCGT	2460
2461	QY	GTGGACATAGTTTGGAAACAAGTCTCAACAGATATCTGAGGATCTTTGTAAGCATATCATC	2520
2461	DB	GTGGACATAGTTTGGAAACAAGTCTCAACAGATATCTGAGGATCTTTGTAAGCATATCATC	2520
2521	QY	TGCCACATCTCTAGGACCTTGTGCTGCTTCATATGTACGTTTTAAGCAGGCTGGGTCTGTTTT	2580
2521	DB	TGCCACATCTCTAGGACCTTGTGCTGCTTCATATGTACGTTTTAAGCAGGCTGGGTCTGTTTT	2580
2581	QY	CAAGACTGCATGAATTTGTGCTTTTCGAGGACTTTCAAGTAGAGGTACTAGTTCGAATATC	2640
2581	DB	CAAGACTGCATGAATTTGTGCTTTTCGAGGACTTTCAAGTAGAGGTACTAGTTCGAATATC	2640
2641	QY	CTTTACGTTGTCTGGTGTGGTGGCCAGGTTGTGCTGAGATGTGGCGAAGAAATGGAC	2700
2641	DB	CTTTACGTTGTCTGGTGTGGTGGCCAGGTTGTGCTGAGATGTGGCGAAGAAATGGAC	2700
2701	QY	TGTCCTCTTATTAGCCAGGTGTTTTATTACCAAGATGTTAAGTCACAGAGAAATCTATG	2760
2701	DB	TGTCCTCTTATTAGCCAGGTGTTTTATTACCAAGATGTTAAGTCACAGAGAAATCTATG	2760
2761	QY	ATAAAGATATCATCATGCTTCAGATTGGTGCATCTTTAATGGAATCCCAATAGTCTTTGT	2820
2761	DB	ATAAAGATATCATCATGCTTCAGATTGGTGCATCTTTAATGGAATCCCAATAGTCTTTGT	2820
2821	QY	TACTGGTACTTCAGAGGTATGAACCTTGCAGGCTTTTAAACAAGCATATCTACAAAG	2880
2821	DB	TACTGGTACTTCAGAGGTATGAACCTTGCAGGCTTTTAAACAAGCATATCTACAAAG	2880

QY	3961	CTGTACTGAAAGGAGTGCTGACGTCCCTTTGCAAGAAACAGAGGTGA	4020
Db	3961		
QY	4021	TAGAAAATATGCCATGGTATATATCGGCTGTGTCAGAAATCTACTGCC	4080
Db	4021		
QY	4081	ACAGGGGAAACCCATAGAACTCTCAGAGAGCCCTAGACCACTTTTCATG	4140
Db	4081		
QY	4141	ACTTGGCATATGGAATCTTATACAGAACTGTGGTCACTGTAATGACGAGT	4200
Db	4141		
QY	4201	AGAAGTATTTGAAGCTGTACAGCTGAGCTCTCAGCAGCGCATTCATG	4260
Db	4201		
QY	4261	ACTTGGAAAGTGAGAAATATCTTTGCCCTCTTTGCAAACTCTGTGCAAT	4320
Db	4261		
QY	4321	CCATTATTTCTTTGCAACTCAAAAGATAAACAGTGAGAAATGCGATGCT	4380
Db	4321		
QY	4381	TTTTGACCCCTGGCAGCGTGATACAGACTGTTCTGGCCAGAAATACAG	4440
Db	4381		
QY	4441	GACATGCTAAAGGAGAAACCAATCTTCTTAATCAAGGAATGGAGATTCT	4500
Db	4441		
QY	4501	CTTTGGAGTTCATTTCCATCTCGAGTCTTTGGGTGAGTCTTCGATTAA	4560
Db	4501		
QY	4561	GCATCAAGGAATGGTATTTCTTTGGCAACAAATTTATAGAAATGGATG	4620
Db	4561		
QY	4621	CACCTGATGAAAGGATCTCGAGTCCCACATGCTGACCTGGAGCACCT	4680
Db	4621		
QY	4681	TCCAGGCAATGAAATCTATTTGGGAGATGAAGAAACCTCTGTTTGGAG	4740
Db	4681		
QY	4741	ATAGGCAGCAATATGGTCTGAAAGCAATTAATGCACTTTGCACTTGC	4800
Db	4741		
QY	4801	GTCTCAGTCTCTGATACAGAAACATCTGGTTGCTCTCTATCAGTTGT	4860
Db	4801		
QY	4861	TAAATCAGAGATACACCATGCTCTGCTATAGATCTGTTTCAATGTTG	4920
Db	4861		
QY	4921	CTGTGTAGCAATCCCATCTTTGATTTGGGATGACCCCTGTTGATCTG	4980
Db	4921		
QY	4981	TTAGTCTCTCTATACCAACCTTTATCTCTCCATTTGATACCAATGGC	5040
Db	4981		
QY	5041	AGATACTACTTACATAGACACAGGCTTACCCCTTGCTCAGGTTCAAGA	5100
Db	5041		

Db	5041	AGATACTACTTACATAGACACAGGCTTACCCCTTGCTCAGGTTCAAGA	5100
QY	5101	AGGTCATTCGGCATCTTCTTTCTTTCAGAAATTTCTCAATATACAGTGG	5160
Db	5101		
QY	5161	GGTGTGATATTTCCCTGGCTGTTTGGGTCTCACATGAAGAAATGGCAT	5220
Db	5161		
QY	5221	TTCCCTGTGTGCAATTTTTCACATATTTTACCTTGGGTAACTCCGCT	5280
Db	5221		
QY	5281	ATACCAATTTCTGCAAGAGGAGTACAGTGCATCTCTAGCTATCTATCT	5340
Db	5281		
QY	5341	ATTTGTTCTGCTCTTCCAGGAATATTTGGGATCTGTAAGGCCCTTGT	5400
Db	5341		
QY	5401	GTGCAGATCTCTGCTTACTAACTGTTTGAAGCAAAAAACACCGTGGT	5460
Db	5401		
QY	5461	GAAAAAGAAATAGTTTGTATAGAGTCTCTGATGATCTATAGTGCCT	5520
Db	5461		
QY	5521	CTCATTTTCAGTGCCACGGTCTGCAGATGATGAGCAAGATCTCTCTCT	5580
Db	5521		
QY	5581	TCTGTGGGCTTACTATGTTCTCAGAACTTTTCTGCCAGGAAATTTGT	5640
Db	5581		
QY	5641	AGTTTGGAGCTTGCAATTTTTCACGCACTTCACTGTGAGCCGGAGT	5700
Db	5641		
QY	5701	AAATCAGAGATGCCGAGTGTCTGTTGAAGTAAAGCCAGAGGCTGTSC	5760
Db	5701		
QY	5761	CTCCTTACTTGGATGAATATGGAGAAACAGACCCCTGGCTGAAGGG	5820
Db	5761		
QY	5821	ATTTATCTGTGAGCGGTATCGGAAGCTCCATTTGGTCTGCAACACAC	5880
Db	5821		
QY	5881	AAGGATTTCTGAGAGCCAGAGACTAATCAGATGTTATTTGGATTCAA	5940
Db	5881		
QY	5941	TGTGAGCTTCAACTCTGCTCAAGCAATCAAAATGACGACGTAGTAAG	6000
Db	5941		
QY	6001	AAAAATATGGAATCTTTCTGAGGCTGGGAAAGTATTTGGAGGCTTT	6060
Db	6001		
QY	6061	CAGGTTCACTTACATCAATAAATATTTCTTAATGGAGTATTTGCTCA	6120
Db	6061		
QY	6121	TATGCTTTCACAGGAAAAAGGACATAGATCAATCTGTTTATGCTAGT	6180
Db	6121		

Db 6121 TATGCTTACAGGAAAGGACATAGATCAATCTGTTTATGTGCTAGTATTTCCAGGA 6180
 QY 6181 ATTATTCCTCCCTTCAATAATTTGCTCAATTTATTTATTCATCCACTTGGTAGATGAAG 6240
 Db 6181 ATTATTCCTCCCTTCAATAATTTGCTCAATTTATTTATTCATCCACTTGGTAGATGAAG 6240
 QY 6241 TCACGTCAACAGTGTAGACATTTTATGTTGGTTAACTCTTCTGCAATTTTGTATTT 6300
 Db 6241 TCACGTCAACAGTGTAGACATTTTATGTTGGTTAACTCTTCTGCAATTTTGTATTT 6300
 QY 6301 GGTGTTTT 6308
 Db 6301 GGTGTTTT 6308

RESULT 2

ADSR6863
 ID ADSR6863 standard; cDNA; 6308 BP.

XX
 AC ADSR6863;

XX
 DT 16-DEC-2004 (first entry)

XX
 DE cDNA encoding human E3alpha ubiquitin ligase, huE3alphaI.

XX
 KW Human; E3alpha ubiquitin ligase; huE3alphaI;
 KW ubiquitin-proteasome pathway; chromosome 15 q; rapid muscle wasting;
 KW fasting; metabolic acidosis; muscle degeneration; kidney failure;
 KW renal cachexia; uremia; diabetes mellitus; sepsis; AIDS wasting syndrome;
 KW cancer cachexia; negative nitrogen balance; burn; Cushing's syndrome;
 KW inflammatory cachexia; hyperthyroidism; denervation atrophy;
 KW protein/tissue wasting; energy-protein malnutrition; muscle atrophy;
 KW gene therapy; ss; gene.

XX
 OS Homo sapiens.

XX
 FH Key Location/Qualifiers

FT 5'UTR 1..695

FT /*tag= a

FT CDS 696..5945

FT /*tag= b

FT /product= "E3alphaI"

FT 5946..6308

FT /*tag= c

XX
 US2004185037-A1.

XX
 PD 23-SEP-2004.

XX
 PF 15-JAN-2004; 2004US-00758672.

XX
 PR 08-MAR-2000; 2000US-0187911P.

XX
 PR 28-NOV-2000; 2000US-00724126.

XX
 XX (HANH/) HAN H.

XX
 PA (KWAK/) KWAK K.

XX
 PI Han H, Kwak K;

XX
 PI Han H, Kwak K;

XX
 DR WPI; 2004-707854/69.

XX
 DR P-P5DB; ADSR6864.

XX
 PT Novel isolated human E3alpha ubiquitin ligase nucleic acid molecule
 PT useful for treating and/or preventing renal cachexia or inflammatory
 PT cachexia.

XX
 PS Claim 1; SEQ ID NO 1; 115pp; English.

XX
 CC The present invention relates to new orthologue of human E3alpha
 CC ubiquitin ligase, huE3alphaI and huE3alphaII. Most intracellular proteins
 CC are degraded through the ubiquitin-proteasome pathway. Proteins are
 CC marked for proteasomal degradation by conjugation of ubiquitin to the
 CC protein. Conjugation of the ubiquitin molecule involves the activation by

CC E1 enzyme, subsequent transfer to E2 enzyme, which serves as a carrier,
 CC and then interacts with a specific E3 ligase family member. E3 ligase
 CC binds to proteins targeted for degradation and catalyses the transfer of
 CC ubiquitin from the E2 carrier enzyme to the target protein. E3 ligase
 CC determines the specificity of the system. The E3alpha family is the main
 CC family of intracellular ligases and is involved in the N-end rule pathway
 CC of protein degradation. E3alpha enzyme binds directly to the primary
 CC destabilising N-terminal amino acid and catalyses ubiquitin conjugation
 CC thereby targeting the protein for degradation. The human E3alpha gene is
 CC located on chromosome 15 q. Increased proteolysis through the ubiquitin-
 CC proteasome pathway has been determined to be a major cause of rapid
 CC muscle wasting including, fasting, metabolic acidosis, muscle
 CC degeneration, kidney failure, renal cachexia, uremia, diabetes mellitus,
 CC sepsis, AIDS wasting syndrome, cancer cachexia, negative nitrogen
 CC balance, burns, Cushing's syndrome, inflammatory cachexia,
 CC hyperthyroidism, denervation atrophy, protein/tissue wasting, and energy-
 CC protein malnutrition. E3alpha plays a role in the overall increase in
 CC ubiquitination that is associated with and may mediate muscle atrophy in
 CC catabolic and other disease states. Treatment may be administered by gene
 CC therapy, cell therapy and antisense therapy methods. The present sequence
 CC is cDNA encoding human E3alphaI.

XX
 SQ Sequence 6308 BP; 1856 A; 1301 C; 1390 G; 1761 T; 0 U; 0 Other;

Query Match 100.0%; Score 6308; DB 13; Length 6308;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 6308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCAAGATTCCGACAGGGGGAAGCTGAGCCAGGACCAATTAATTTACTTTTACCT 60

Db 1 GCCAAGATTCCGACAGGGGGAAGCTGAGCCAGGACCAATTAATTTACTTTTACCT 60

QY 61 CATTTGTGTAAGACAAAGCGTCAAAAACAGCTTCAACCTATCTTGAACAAGAACTTACCT 120

Db 61 CATTTGTGTAAGACAAAGCGTCAAAAACAGCTTCAACCTATCTTGAACAAGAACTTACCT 120

QY 121 CCAAGGCTTATCATCTGTCTTCCACTTATCCAAAGCTGTATGGCCACTGCTGTGC 180

Db 121 CCAAGGCTTATCATCTGTCTTCCACTTATCCAAAGCTGTATGGCCACTGCTGTGC 180

QY 181 CGCAGCTGGAAACACCGCCAGCCCACTACTGCGCTCCACTACCTGCTGTTCTCCACCT 240

Db 181 CGCAGCTGGAAACACCGCCAGCCCACTACTGCGCTCCACTACCTGCTGTTCTCCACCT 240

QY 241 GATCAGCTGCTGTGCTGCCATCTTATCCGCTTCTGCTGCTTCTGAGTAAATGTATACA 300

Db 241 GATCAGCTGCTGTGCTGCCATCTTATCCGCTTCTGCTGCTTCTGAGTAAATGTATACA 300

QY 301 CACCTGGAAACACCGCCAGCCCACTACTGCGCTCCACTACCTGCTGTTCTCCACCT 360

Db 301 CACCTGGAAACACCGCCAGCCCACTACTGCGCTCCACTACCTGCTGTTCTCCACCT 360

QY 361 CCGAGCTTTTGTACACATGTGCAAGTCCCAAGGCTAGATCTCTAAATAGAGAT 420

Db 361 CCGAGCTTTTGTACACATGTGCAAGTCCCAAGGCTAGATCTCTAAATAGAGAT 420

QY 421 GTATGCAACAGTTCCTCAGCACAAACCCAGATATACACCTTACAGTACCAAGAGTAC 480

Db 421 GTATGCAACAGTTCCTCAGCACAAACCCAGATATACACCTTACAGTACCAAGAGTAC 480

QY 481 GCCTGATAAATTAGAGGGGAAAAAATCTCCAGTCCCTTTCAGCTCGTGAAGCTTGC 540

Db 481 GCCTGATAAATTAGAGGGGAAAAAATCTCCAGTCCCTTTCAGCTCGTGAAGCTTGC 540

QY 541 TCCGGAGAGCGGCGCGGAGCCACTCTTCGAGTCTGGGTCAACCCGAGCTTCAGGGGCG 600

Db 541 TCCGGAGAGCGGCGCGGAGCCACTCTTCGAGTCTGGGTCAACCCGAGCTTCAGGGGCG 600

QY 601 TCGTAAAGTCTGCTCCCTGCTCTCCGACCGGACACAGGTTTCGCTTGTCTGTCGCG 660

Db 601 TCGTAAAGTCTGCTCCCTGCTCTCCGACCGGACACAGGTTTCGCTTGTCTGTCGCG 660

QY 661 GGGGTGCGCAACTGCGAGCGCTGAGTTTCCCTCAAGATGCGGAGCGAGGCTGAGGTA 720

Db 661 ||||| GGGGTCGCGCACTGCGAGGTCAGTTTCCCTCAGATGCGGACGAGGAGGCTGGAGTA 720
Qy 721 CTGAGAGATGGAATTCAGCGCGAGTTTACCCAGACCCCTCAGCGTCTGGCATCTTGGT 780
Db 721 CTGAGAGATGGAATTCAGCGCGAGTTTACCCAGACCCCTCAGCGTCTGGCATCTTGGT 780
Qy 781 GGGATCAGCAAGTTGATTTTATATCTGCTTTCTTCGATCATTTGGCAAAATTTGGTGCCAG 840
Db 781 GGGATCAGCAAGTTGATTTTATATCTGCTTTCTTCGATCATTTGGCAAAATTTGGTGCCAG 840
Qy 841 AAATTTACTTTGCTGGAATTCAGACCTTGGAAAAGCAGGAGAAAGTGTACAAATGT 900
Db 841 AAATTTACTTTGCTGGAATTCAGACCTTGGAAAAGCAGGAGAAAGTGTACAAATGT 900
Qy 901 CAATATTCATCTCACTGGAATTCATTTATTTGGAGAAAGTCCAGATATTTGCTTAGAGA 960
Db 901 CAATATTCATCTCACTGGAATTCATTTATTTGGAGAAAGTCCAGATATTTGCTTAGAGA 960
Qy 961 AATTGAAGCACAGTGGAGCAATTCAGCTTTGTTGGAGGGTTTCAAAGTGGAGAGACAA 1020
Db 961 AATTGAAGCACAGTGGAGCAATTCAGCTTTGTTGGAGGGTTTCAAAGTGGAGAGACAA 1020
Qy 1021 CCTATTTCTGAGGATTTGTCAAATTTGATCCAAATGTTGATCTCTGTATGGAATCTCTCC 1080
Db 1021 CCTATTTCTGAGGATTTGTCAAATTTGATCCAAATGTTGATCTCTGTATGGAATCTCTCC 1080
Qy 1081 AGGACAGTGTTCATAAAATCATCTGTTACAGATGCATATTTCTACTGGAGAGGGTTCT 1140
Db 1081 AGGACAGTGTTCATAAAATCATCTGTTACAGATGCATATTTCTACTGGAGAGGGTTCT 1140
Qy 1141 GTGACTGTGGAGACACAGAGGATTCGAAAACCTGGCCCTTTTGTGTAAATCATGAACCTG 1200
Db 1141 GTGACTGTGGAGACACAGAGGATTCGAAAACCTGGCCCTTTTGTGTAAATCATGAACCTG 1200
Qy 1201 GAAGAGCAGGTACTATAAAGAGAAATTCACGCTGTCGTTGAATGAAGAGGTAAATTTGTC 1260
Db 1201 GAAGAGCAGGTACTATAAAGAGAAATTCACGCTGTCGTTGAATGAAGAGGTAAATTTGTC 1260
Qy 1261 AAGCCAGGAAAATATTTCTTCAGTGTATAAAATATGTCGTAGAAATGATATATCGGAAG 1320
Db 1261 AAGCCAGGAAAATATTTCTTCAGTGTATAAAATATGTCGTAGAAATGATATATCGGAAG 1320
Qy 1321 AGGAAAAGAACTGCTCTGAACTCCAGATTAAGGAGAAAATGAAAGATATTTGTC 1380
Db 1321 AGGAAAAGAACTGCTCTGAACTCCAGATTAAGGAGAAAATGAAAGATATTTGTC 1380
Qy 1381 TCTTTTCAATGATGAACACCAATTCATATGACCACTCATATACAGCCCTACAAAGAGCTC 1440
Db 1381 TCTTTTCAATGATGAACACCAATTCATATGACCACTCATATACAGCCCTACAAAGAGCTC 1440
Qy 1441 TTGACTGTGAGCTCGCAGAGGCCCATACCACTGTCATACCACTGTCATACAAAGAGGTCGTC 1500
Db 1441 TTGACTGTGAGCTCGCAGAGGCCCATACCACTGTCATACCACTGTCATACAAAGAGGTCGTC 1500
Qy 1501 GGGCTGTTAAAGCGGAGCTTATGCTTGGCCAGGACCAAGCAAGATATAAAGAGCTC 1560
Db 1501 GGGCTGTTAAAGCGGAGCTTATGCTTGGCCAGGACCAAGCAAGATATAAAGAGCTC 1560
Qy 1561 ATTCAGAAAATGCTCTCAACATCCACTTCATGTAGAGTATTAACAATCAGAGATTTATGG 1620
Db 1561 ATTCAGAAAATGCTCTCAACATCCACTTCATGTAGAGTATTAACAATCAGAGATTTATGG 1620
Qy 1621 CTTCATCAGAAAATTTGCTTTGCGTCTTGGTTCTCGATGAACAAAATTTATGACTATTCAA 1680
Db 1621 CTTCATCAGAAAATTTGCTTTGCGTCTTGGTTCTCGATGAACAAAATTTATGACTATTCAA 1680
Qy 1681 GTGACTTTAGGAGATCTTTTGGCAAGCATGCTTAGAGAGAACTTGACTCGGAGATC 1740
Db 1681 GTGACTTTAGGAGATCTTTTGGCAAGCATGCTTAGAGAGAACTTGACTCGGAGATC 1740
Qy 1741 CTGTCTCTAAGCAGGTAAATGCTTTGGGATGCAAGCTTTATAAAGGTGCCCGCTAAGA 1800

Db 1741 CCTCTCTATAAGCAGGTTAAATGCTTTGGGATGCAAAAGCTTTATAAAAGTGCCCGTAAGA 1800
Qy 1801 TCCTTCATGAATTCATCTTCAGCAGTTTATTTTATGGAGATGGAATACAAAACCTCTTTG 1860
Db 1801 TCCTTCATGAATTCATCTTCAGCAGTTTATTTTATGGAGATGGAATACAAAACCTCTTTG 1860
Qy 1861 CTATGGAATTTGTGAAGTATTTATAACAACCTGCGAGAAAGATATATCAGTGATGATCATG 1920
Db 1861 CTATGGAATTTGTGAAGTATTTATAACAACCTGCGAGAAAGATATATCAGTGATGATCATG 1920
Qy 1921 ACAGAGTATCTCTATAAATCTGCACTTTCACTTCAGATGATTTTACTGTTCTTACTCTGGCTC 1980
Db 1921 ACAGAGTATCTCTATAAATCTGCACTTTCACTTCAGATGATTTTACTGTTCTTACTCTGGCTC 1980
Qy 1981 GACATCTTATTTGAAGAGCAGAAATGTTATCTCTGTCATTTACTGAAACTCTGCTAGAGTTT 2040
Db 1981 GACATCTTATTTGAAGAGCAGAAATGTTATCTCTGTCATTTACTGAAACTCTGCTAGAGTTT 2040
Qy 2041 TACCTGAGTACTTTCGACAGGAAACAAATAATTTCAACTTCCAGGGTTATAGCCAGGACAAAT 2100
Db 2041 TACCTGAGTACTTTCGACAGGAAACAAATAATTTCAACTTCCAGGGTTATAGCCAGGACAAAT 2100
Qy 2101 TGGAGAGATATATGCGAGTAAATATGTCACCTTAAAGTATATCTCTGATCAGCAAAACCCCAA 2160
Db 2101 TGGAGAGATATATGCGAGTAAATATGTCACCTTAAAGTATATCTCTGATCAGCAAAACCCCAA 2160
Qy 2161 TATGGAAGAGATTTAAGATGAGTTCCTTTGAAGTTTTCGATCTTTTGAAGATTC 2220
Db 2161 TATGGAAGAGATTTAAGATGAGTTCCTTTGAAGTTTTCGATCTTTTGAAGATTC 2220
Qy 2221 TTACTGTATGTCAGGGAATGGAAGAAATCCGAAGACAGTTGGGCAACACATTTGAAGTGG 2280
Db 2221 TTACTGTATGTCAGGGAATGGAAGAAATCCGAAGACAGTTGGGCAACACATTTGAAGTGG 2280
Qy 2281 ATCTGATTTGGAGGGTCCATTCATATACAGATGCAATTAAGAAATTTTATCTCATGT 2340
Db 2281 ATCTGATTTGGAGGGTCCATTCATATACAGATGCAATTAAGAAATTTTATCTCATGT 2340
Qy 2341 TCCAAGATGCTGCTGTCATGATGAAGAACTTCTTACTGTCGCTTATAAAGATGTCACA 2400
Db 2341 TCCAAGATGCTGCTGTCATGATGAAGAACTTCTTACTGTCGCTTATAAAGATGTCACA 2400
Qy 2401 AAGCTGTGATCAGGTGCACTACAGTTTTCATATCTAGTAGCAAGACAGTAGTACAATCGT 2460
Db 2401 AAGCTGTGATCAGGTGCACTACAGTTTTCATATCTAGTAGCAAGACAGTAGTACAATCGT 2460
Qy 2461 GTGACATAGTTTGGAAAACAAAGTCCACAGATATCTGAGGATCTTGTAGGACATACATC 2520
Db 2461 GTGACATAGTTTGGAAAACAAAGTCCACAGATATCTGAGGATCTTGTAGGACATACATC 2520
Qy 2521 TGCCACTCTCTAGGACCTTCTGCTGTCATGTCATGTTAAGCAGGCTGGGTCTGTTT 2580
Db 2521 TGCCACTCTCTAGGACCTTCTGCTGTCATGTCATGTTAAGCAGGCTGGGTCTGTTT 2580
Qy 2581 CAAGACTGCATGAATTTGTGCTTTTGAAGGACTTTCAAGTAGAGGTACTAGTGGAAATATC 2640
Db 2581 CAAGACTGCATGAATTTGTGCTTTTGAAGGACTTTCAAGTAGAGGTACTAGTGGAAATATC 2640
Qy 2641 CTTTACGTTGTCTGGTGTGGTGGTCCAGGTTGTTGCTGAGATGTTGGCGAAGAAATGGAC 2700
Db 2641 CTTTACGTTGTCTGGTGTGGTGGTCCAGGTTGTTGCTGAGATGTTGGCGAAGAAATGGAC 2700
Qy 2701 TGTCTCTTATTTAGCCAGGTTTATTTACCAAGATGTTAAGTGCAGAGAAATGATATG 2760
Db 2701 TGTCTCTTATTTAGCCAGGTTTATTTACCAAGATGTTAAGTGCAGAGAAATGATATG 2760
Qy 2761 ATAAAGATATCATCATGCTTCAGATTTGTCATCTTTAATGGATCCCAATTAAGTTCTTGT 2820
Db 2761 ATAAAGATATCATCATGCTTCAGATTTGTCATCTTTAATGGATCCCAATTAAGTTCTTGT 2820
Qy 2821 TACTGGTACTTTCAGAGGTATGAATTCGCGAGGCTTTTAAACAGACCATATCTACAAAG 2880
Db 2821 TACTGGTACTTTCAGAGGTATGAATTCGCGAGGCTTTTAAACAGACCATATCTACAAAG 2880

QY 2881 ACCAGGATTTGATTTAAACAAATATATATACATTAATAGAAAGAAATGCTTCAGGTCCTCATCT 2940
DB 2881 ACCAGGATTTGATTTAAACAATATATATACATTAATAGAAAGAAATGCTTCAGGTCCTCATCT 2940
QY 2941 ATATTGTCGGGTGAGCGTTATGTACCTGAGGTGGGAAATGTGACCAAAAGAGGTCACAA 3000
DB 2941 ATATTGTCGGGTGAGCGTTATGTACCTGAGGTGGGAAATGTGACCAAAAGAGGTCACAA 3000
QY 3001 TGAGAGAAATCAATTCACATTCGCTTTGTCATTGAAACCCATGCCACACAGTGCCTATGCAAAA 3060
DB 3001 TGAGAGAAATCAATTCACATTCGCTTTGTCATTGAAACCCATGCCACACAGTGCCTATGCAAAA 3060
QY 3061 ATTTACCTGAGAAATGAATAATATGAAACTGCGCTTAGAGAAATGTATATAACAAAGTGGCCA 3120
DB 3061 ATTTACCTGAGAAATGAATAATATGAAACTGCGCTTAGAGAAATGTATATAACAAAGTGGCCA 3120
QY 3121 CATTTTAAAGAACCCAGGTGTATCAGGCCATGAGTTTATGAACTAAAGATGAATCACTGA 3180
DB 3121 CATTTTAAAGAACCCAGGTGTATCAGGCCATGAGTTTATGAACTAAAGATGAATCACTGA 3180
QY 3181 AAGACTTCAATATGTACTTTTATCAATTAATCACTTCAAAACCCAGCATAGCAAGCTGAACATA 3240
DB 3181 AAGACTTCAATATGTACTTTTATCAATTAATCACTTCAAAACCCAGCATAGCAAGCTGAACATA 3240
QY 3241 TGCAGAGAAAGAGGAGAAACAAAGAAACAAAGATGAAGCATTTGCGGCCACACACCTC 3300
DB 3241 TGCAGAGAAAGAGGAGAAACAAAGAAACAAAGATGAAGCATTTGCGGCCACACACCTC 3300
QY 3301 CTGAATTCGCGCTCTTTCAGCAAGGTGATTAACCTTCTCACTGTGATATCATGATGT 3360
DB 3301 CTGAATTCGCGCTCTTTCAGCAAGGTGATTAACCTTCTCACTGTGATATCATGATGT 3360
QY 3361 ACATTTCTCAGAACCGTATTTGAGCGGGCAATAGACACAGATTTCTAACTTTGTGGACCGAAG 3420
DB 3361 ACATTTCTCAGAACCGTATTTGAGCGGGCAATAGACACAGATTTCTAACTTTGTGGACCGAAG 3420
QY 3421 GGATGCTCCAAATGGCTTTTATATTTCTGGCAATTTGGGTTTACTAGAGAGAACCAACAGC 3480
DB 3421 GGATGCTCCAAATGGCTTTTATATTTCTGGCAATTTGGGTTTACTAGAGAGAACCAACAGC 3480
QY 3481 TTCAAAAAGCTCTCTGAGAGAGATTAACATTTGACTTTTATCATAGAGCTTCAAGATTGG 3540
DB 3481 TTCAAAAAGCTCTCTGAGAGAGATTAACATTTGACTTTTATCATAGAGCTTCAAGATTGG 3540
QY 3541 GAAGTTTCAAGCATGAATATACAAATGCTTTTGGAAAACTCAAGGAATTTCCCGATTAG 3600
DB 3541 GAAGTTTCAAGCATGAATATACAAATGCTTTTGGAAAACTCAAGGAATTTCCCGATTAG 3600
QY 3601 AAGGCCAGAGGACATGATAACGTGGATACCTTCAGATGTTTGAACAGTGAAGCGATTAA 3660
DB 3601 AAGGCCAGAGGACATGATAACGTGGATACCTTCAGATGTTTGAACAGTGAAGCGATTAA 3660
QY 3661 GAGAAAAATCTTGTTTAAATTTAGCAACACATCAGGATCGGAATCTATTAAGAATGATG 3720
DB 3661 GAGAAAAATCTTGTTTAAATTTAGCAACACATCAGGATCGGAATCTATTAAGAATGATG 3720
QY 3721 AGATTTACTCATGATAAAGAAAGCAGAACGAAAGAAAGAGCTGAAGCTCTAGGCTAC 3780
DB 3721 AGATTTACTCATGATAAAGAAAGCAGAACGAAAGAAAGAGCTGAAGCTCTAGGCTAC 3780
QY 3781 ATCGCCAGAGATCATGCTCAGATGCTGCTGCTTACAGAAAACTTCAATTTGAAACTCAT 3840
DB 3781 ATCGCCAGAGATCATGCTCAGATGCTGCTGCTTACAGAAAACTTCAATTTGAAACTCAT 3840
QY 3841 AACTCATGATGACAAATCATCAGAAATGCTGGGAAAGAGATTTCCATTTATGAGGGAAG 3900
DB 3841 AACTCATGATGACAAATCATCAGAAATGCTGGGAAAGAGATTTCCATTTATGAGGGAAG 3900
QY 3901 AGAGACCCAGCAGTCACTGCTACTAGATTTGCTTTGGTCTTAAACGGGCTCCAT 3960
DB 3901 AGAGACCCAGCAGTCACTGCTACTAGATTTGCTTTGGTCTTAAACGGGCTCCAT 3960

QY 3961 CTGTTACTGAAAAAGGAGTGTCTGACGTGCATCCTTTGCCAAGAAAGACAGAGGTGAAAA 4020
DB 3961 CTGTTACTGAAAAAGGAGTGTCTGACGTGCATCCTTTGCCAAGAAAGACAGAGGTGAAAA 4020
QY 4021 TAGAAAAATAATGCAATGTGATTTATCGGCTCTGTGTCAGAAATCTACTGCCCTTAACCCAGC 4080
DB 4021 TAGAAAAATAATGCAATGTGATTTATCGGCTCTGTGTCAGAAATCTACTGCCCTTAACCCAGC 4080
QY 4081 ACAGGGGAAAAACCATAGAACTCTCAGGAGAAAGCCCTTAGACCCACTTTTTCATGGATCCAG 4140
DB 4081 ACAGGGGAAAAACCATAGAACTCTCAGGAGAAAGCCCTTAGACCCACTTTTTCATGGATCCAG 4140
QY 4141 ACTTTGGCATATGGAACCTTATACAGAGAGCTGTGCTCATGTAATGCAAGCAGTGTCTGGC 4200
DB 4141 ACTTTGGCATATGGAACCTTATACAGAGAGCTGTGCTCATGTAATGCAAGCAGTGTCTGGC 4200
QY 4201 AGAAGTATTTTGAAGCTGTATACAGCTGAGCTCTCAGCAGCGCATTCATGTCACCTTTTGTG 4260
DB 4201 AGAAGTATTTTGAAGCTGTATACAGCTGAGCTCTCAGCAGCGCATTCATGTCACCTTTTGTG 4260
QY 4261 ACTTTGGAAGTGGAGAAATATCTTTTCCCTCTTTTGCCTCTCTGTGCAATCTCTGTGCAATCTCTGTGATCC 4320
DB 4261 ACTTTGGAAGTGGAGAAATATCTTTTCCCTCTTTTGCCTCTCTGTGCAATCTCTGTGATCC 4320
QY 4321 CCATTTATCTTTTGCACCTCAAAAGATAAACAAGTGAATGAGATGCTCTTGTCTCAAC 4380
DB 4321 CCATTTATCTTTTGCACCTCAAAAGATAAACAAGTGAATGAGATGCTCTTGTCTCAAC 4380
QY 4381 TTTTGGCCCTGGCAGCGTGGATACAGACTGTTCTGGCCAGAAATATCAGGTTTATATATAA 4440
DB 4381 TTTTGGCCCTGGCAGCGTGGATACAGACTGTTCTGGCCAGAAATATCAGGTTTATATATAA 4440
QY 4441 GACATGCTTAAAGGAGAAAAACCAATTCCTATTTTCTTTTAAATCAAGGATGGAGATTTCTA 4500
DB 4441 GACATGCTTAAAGGAGAAAAACCAATTCCTATTTTCTTTTAAATCAAGGATGGAGATTTCTA 4500
QY 4501 CTTTGGAGTTCATTCATCTGAGTTTGGGCTTGGTGTGAGTCTTCGATTAATTTCAAAATA 4560
DB 4501 CTTTGGAGTTCATTCATCTGAGTTTGGGCTTGGTGTGAGTCTTCGATTAATTTCAAAATA 4560
QY 4561 GCATCAAGGAAATGTTTATCTCTTTCGCAACAAATTTATAGAAATGGATTTGAAAGTGC 4620
DB 4561 GCATCAAGGAAATGTTTATCTCTTTCGCAACAAATTTATAGAAATGGATTTGAAAGTGC 4620
QY 4621 CACCTGATGAAGGAGTCTCGAGTCCCATCTGCTGAGTCCCATCTGCTGAGCAGCCTTCACTA 4680
DB 4621 CACCTGATGAAGGAGTCTCGAGTCCCATCTGCTGAGTCCCATCTGCTGAGCAGCCTTCACTA 4680
QY 4681 TCCAGGCAATGAAAAATCTATTTGGGAGATGAAGAAAAACCTCTGTTTGGAGCATTTCAAA 4740
DB 4681 TCCAGGCAATGAAAAATCTATTTGGGAGATGAAGAAAAACCTCTGTTTGGAGCATTTCAAA 4740
QY 4741 ATAGGAGCATTAATGCTGCAAGCATTTAATGCAAGTTTGGAGTTGCAAGAGGATTAACCT 4800
DB 4741 ATAGGAGCATTAATGCTGCAAGCATTTAATGCAAGTTTGGAGTTGCAAGAGGATTAACCT 4800
QY 4801 GTCCTCAGGTCCTCATATACAGAAACATCTGGTTCGCTCTTCTATCAGTTGTTCTTCTCAACA 4860
DB 4801 GTCCTCAGGTCCTCATATACAGAAACATCTGGTTCGCTCTTCTATCAGTTGTTCTTCTCAACA 4860
QY 4861 TAAATCAGAAAGATACACCATGCTTCTGTCTATAGATCTGTTTTCATGTTTGGTGGTG 4920
DB 4861 TAAATCAGAAAGATACACCATGCTTCTGTCTATAGATCTGTTTTCATGTTTGGTGGTG 4920
QY 4921 CTGTTTAGCATTTCCCATCTTGTATTTGGGATGACCTGTTGATCTGAGCCTTCTTCAG 4980
DB 4921 CTGTTTAGCATTTCCCATCTTGTATTTGGGATGACCTGTTGATCTGAGCCTTCTTCAG 4980
QY 4981 TTAGTCTTCTCTATAACCACTTTTATCTTCCATTTTGTATCCATGCAACCATGCTTC 5040
DB 4981 TTAGTCTTCTCTATAACCACTTTTATCTTCCATTTTGTATGCAACCATGCTTC 5040
QY 5041 AGATATCTTACAGTATAGACACAGGCTTACCCCTTGTCTCAGGTTCAAGAGACAGTGAAG 5100

Db 5041 AGATACTACTTACAGTACAGACAGGCGCTACCCCTTGCTCAGGTTCAAGAGACAGTGAAG 5100
Qy 5101 AGGCTCATTCGCGCATCTCTTCTTTGTCAGAGAAATTTCTCAATATACAAAGTGGCTCCATTG 5160
Db 5101 AGGCTCATTCGCGCATCTCTTCTTTGTCAGAGAAATTTCTCAATATACAAAGTGGCTCCATTG 5160
Qy 5161 GGTGTGATATTCCTGGCTGGTATTTGTCGGTCTCACTGAAGAAATGGCAATACCCCTTATC 5220
Db 5161 GGTGTGATATTCCTGGCTGGTATTTGTCGGTCTCACTGAAGAAATGGCAATACCCCTTATC 5220
Qy 5221 TTGCTGTGCTGCAATGTTTTTCACTATTTTCACTTTGCGGTAACTCCGCTCAGGAATGTC 5280
Db 5221 TTGCTGTGCTGCAATGTTTTTCACTATTTTCACTTTGCGGTAACTCCGCTCAGGAATGTC 5280
Qy 5281 ATACCAATTCGAGAGAGAGTACAGTGCACTCTGTAGCTATCTATCTTTTACCTACAA 5340
Db 5281 ATACCAATTCGAGAGAGAGTACAGTGCACTCTGTAGCTATCTATCTTTTACCTACAA 5340
Qy 5341 ATTTGTTCTCTCTTCCAGGAATATGGGATCTGTAAAGGCCCTTGCTCCAGAGGTGGT 5400
Db 5341 ATTTGTTCTCTCTTCCAGGAATATGGGATCTGTAAAGGCCCTTGCTCCAGAGGTGGT 5400
Qy 5401 GTGCAGATCTGCTCTTAACTCTTTTGAAGCAAAAAAACAACCGTGTCAAGTACCCCTA 5460
Db 5401 GTGCAGATCTGCTCTTAACTCTTTTGAAGCAAAAAAACAACCGTGTCAAGTACCCCTA 5460
Qy 5461 GAAAGAAATAGTTGTATAGAGTCTCTGTATGATATAGTGTCTCTGATCAAGCTT 5520
Db 5461 GAAAGAAATAGTTGTATAGAGTCTCTGTATGATATAGTGTCTCTGATCAAGCTT 5520
Qy 5521 CTATTTTCAGGTGCCAGCTCTGAGATGATGAGCGAAGCATCTGCTCTGCTCTT 5580
Db 5521 CTATTTTCAGGTGCCAGCTCTGAGATGATGAGCGAAGCATCTGCTCTGCTCTT 5580
Qy 5581 TCTGTGGGCTATATATGTTCTCAGAAATTTGCTGCGAGAAATTTGTAACGGGGAAG 5640
Db 5581 TCTGTGGGCTATATATGTTCTCAGAAATTTGCTGCGAGAAATTTGTAACGGGGAAG 5640
Qy 5641 AGTTTGGAGCTTGCAATTTTTCAGCACTTCACTGTGAGCGCGAGTGTGCAATTTTCTTAA 5700
Db 5641 AGTTTGGAGCTTGCAATTTTTCAGCACTTCACTGTGAGCGCGAGTGTGCAATTTTCTTAA 5700
Qy 5701 AAATCAGAGATGCGAGTGTCTGCTGTTGAAGTAAAGCCAGAGCTGTGCTATCCAG 5760
Db 5701 AAATCAGAGATGCGAGTGTCTGCTGTTGAAGTAAAGCCAGAGCTGTGCTATCCAG 5760
Qy 5761 CTCTTACTTGGATGAATATGAGAGAAACAGACCTCTGGCTGGAAGGGGCAACCCCTTC 5820
Db 5761 CTCTTACTTGGATGAATATGAGAGAAACAGACCTCTGGCTGGAAGGGGCAACCCCTTC 5820
Qy 5821 ATTTATCTGTCGAGCGGTATCGGAAGCTCCATTTGGTCTGGCAACAACACTGCAATATAG 5880
Db 5821 ATTTATCTGTCGAGCGGTATCGGAAGCTCCATTTGGTCTGGCAACAACACTGCAATATAG 5880
Qy 5881 AAGAGATTTCTAGGAGCCAGAGACTATCAGATGTTATTTGGATTCACTGCGCAGTTAC 5940
Db 5881 AAGAGATTTCTAGGAGCCAGAGACTATCAGATGTTATTTGGATTCACTGCGCAGTTAC 5940
Qy 5941 TGTGAGCTCCAACTCTGCTCAAGCAATCAAAATGACGACAGTAGTAAAGGCTGATTC 6000
Db 5941 TGTGAGCTCCAACTCTGCTCAAGCAATCAAAATGACGACAGTAGTAAAGGCTGATTC 6000
Qy 6001 AAAATATGGAATACTTTCTGAGGCTGGGAAAGTATTTGAGGGTCTTTTGTCTCATGTC 6060
Db 6001 AAAATATGGAATACTTTCTGAGGCTGGGAAAGTATTTGAGGGTCTTTTGTCTCATGTC 6060
Qy 6061 CAGGTTCTACTACATCAATTAATTTCTTAATGAGTATTTGCTTCAATAGCAACA 6120
Db 6061 CAGGTTCTACTACATCAATTAATTTCTTAATGAGTATTTGCTTCAATAGCAACA 6120
Qy 6121 TATGCTTCCAGGAAAAAGGACATAGATCAATCTGTTTATGCTAGTATTTCCAGGA 6180

Db 6121 TATGCTTCCAGGAAAAAGGACATAGATCAATCTGTTTATGCTAGTATTTCCAGGA 6180
Qy 6181 ATTTATTTCCCTTCATATTTGCTCATTTTATTTTATTTTATTCATCCACTTGTAGATGAAG 6240
Db 6181 ATTTATTTCCCTTCATATTTGCTCATTTTATTTTATTTTATTCATCCACTTGTAGATGAAG 6240
Qy 6241 TCAGCTCAAAACAGTTGTAGACATTTTATGTTGTTGTTTAACTCTTCTGCAATTTTGTATTT 6300
Db 6241 TCAGCTCAAAACAGTTGTAGACATTTTATGTTGTTGTTTAACTCTTCTGCAATTTTGTATTT 6300
Qy 6301 GGTGTTTT 6308
Db 6301 GGTGTTTT 6308
RESULT 3
ID AB224689 standard; cDNA; 7742 BP.
XX AC AB224689;
XX DT 07-APR-2003 (first entry)
XX Human cell growth, differentiation and death protein CGDD-1 cDNA.
XX CGDD-1; cell growth; cell differentiation; cell death; human; cytostatic;
XX antiarteriosclerotic; hepatotropic; antiinflammatory; antipsoriatic;
XX antianaemic; ophthalmological; auditory; anticonvulsant;
XX cerebroprotective; nootropic; neuroprotective; anti-HIV; antiallergic;
XX neuroleptic; tranquilizer; immunosuppressive; anti-HIV; antiallergic;
XX antiasthmatic; antithyroid; antidiabetic; dermatological; nephrotropic;
XX antirheumatic; antithyroid; antidiabetic; antidiabetic; virucide;
XX antibacterial; fungicide; antiparasitic; protozoicide; antihelminthic;
XX antiinfertility; gynaecological; ubiquitin protein ligase; enzyme;
XX gene therapy; microarray; gene; ss.
XX Homo sapiens.
XX OS
XX FH Key Location/Qualifiers
XX CDS 13..5229
XX FT /*tag= a
XX FT /product= "Human CGDD-1"
XX W0200297032-A2.
XX PD 05-DEC-2002.
XX PF 05-APR-2002; 2002WO-US011152.
XX PR 06-APR-2001; 2001US-0282110P.
XX PR 11-APR-2001; 2001US-0283294P.
XX PR 26-APR-2001; 2001US-0286820P.
XX PR 27-APR-2001; 2001US-0287228P.
XX PR 16-MAY-2001; 2001US-0291662P.
XX PR 18-MAY-2001; 2001US-0291846P.
XX PR 25-MAY-2001; 2001US-0293727P.
XX PR 01-JUN-2001; 2001US-0295263P.
XX PR 01-JUN-2001; 2001US-0295340P.
XX PR 15-JAN-2002; 2002US-0349705P.
XX (INCY-) INCYTE GENOMICS INC.
XX PA
XX Azimzai Y, Au-Young JK, Batra S, Baughn MR, Becha SD,
XX Borowsky ML, Burford N, Ding L, Elliott VS, Emerling BM, Gandhi AR;
XX Gietzen KU, Griffin JA, Hafalia AJA, Honchell CD, Lal PG, Lee SY;
XX Lu DAM, Arvizu CS, Ramkumar J, Reddy R, Sanjanwala MM, Tang YT;
XX Walia NK, Wang YE, Warren BA, Xu Y, Yang J, Yao MG, Yue H;
XX Zebarjadian Y;
XX WPI; 2003-140453/13.
XX P-PSDB; ABP58330.
XX Novel human proteins associated with cell growth, differentiation and

PT death, useful for treating, diagnosing or preventing cancer,
PT developmental, neurological, reproductive or autoimmune/inflammatory
XX disorders.

PS Claim 5; Page 216-218; 238pp; English.

CC The present sequence is that of Incyte clone 1351608C81 encoding human
CC CGD-1, a novel protein associated with cell growth, differentiation and
CC death. A representative cDNA library for the polynucleotide is P6ANN001
CC from paragonomic tumour tissue. Structural features establish the
CC encoded protein as being associated with cell growth, differentiation and
CC death, with further evidence suggesting it to be a ubiquitin protein
CC ligase. The invention is based on novel human CGD-1 to -21 proteins (see
CC ABP5830-50), the polynucleotides encoding them (see ABZ24689-709), and
CC to the use of these for the diagnosis, treatment or prevention of cell
CC proliferative disorders including cancer, developmental disorders,
CC neurological disorders, autoimmune disorders, reproductive disorders, and
CC disorders of the placenta, and in the assessment of the effects of
CC exogenous compounds on the activity and expression of proteins and
CC nucleic acids associated with cell growth, differentiation and death.
CC CGD polynucleotides are also used in a claimed microarray and in a
CC claimed method of generating an expression profile of a sample

XX SQ Sequence 7742 BP; 2388 A; 1423 C; 1622 G; 2309 T; 0 U; 0 Other;

Query Match 88.6%; Score 5592; DB 8; Length 7742;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5592; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 717 GGTACTGAGGATGGAATCAGCGGAGTTACCCAGACCCCTCAGGCTGCGATCT 776
DB 1 GGTACTGAGGATGGAATCAGCGGAGTTACCCAGACCCCTCAGGCTGCGATCT 60

QY 777 TGGTGGGATCAGCAAGTTGATTTTATCTGCTTTCTTGATCATTTGGCACAATTTGGTG 836
DB 61 TGGTGGGATCAGCAAGTTGATTTTATCTGCTTTCTTGATCATTTGGCACAATTTGGTG 120

QY 837 CCAGAAATTTACTTTGCTGAAATGACCCAGACTTTGGAAAGCAGGAGAAAGTGTACAA 896
DB 121 CCAGAAATTTACTTTGCTGAAATGACCCAGACTTTGGAAAGCAGGAGAAAGTGTACAA 180

QY 897 ATGTCAATTTACTTCCACTGGAATGGTACTTTATTTGGAGAGATCCAGATATTTGCTTA 956
DB 181 ATGTCAATTTACTTCCACTGGAATGGTACTTTATTTGGAGAGATCCAGATATTTGCTTA 240

QY 957 GAGAAATTTGAGCAGTGGAGCATTTTCAGCTTTGCGAGGGTTTCAAAGTGGAGAG 1016
DB 241 GAGAAATTTGAGCAGTGGAGCATTTTCAGCTTTGCGAGGGTTTCAAAGTGGAGAG 300

QY 1017 ACAACCTATTCTTGAGGGATTTGCAATTTGATCCAAATGTAATCTCTGTATGGACTGC 1076
DB 301 ACAACCTATTCTTGAGGGATTTGCAATTTGATCCAAATGTAATCTCTGTATGGACTGC 360

QY 1077 TTCAGGACAGTGTTCATAAAATCATCGTTACAGATGCACTTCTACTGGAGAGGG 1136
DB 361 TTCAGGACAGTGTTCATAAAATCATCGTTACAGATGCACTTCTACTGGAGAGGG 420

QY 1137 TTCTGTGACTGTGGAGACACAGAGCATGGAAATCTGGCCCTTTTGTGTAAATCATGAA 1196
DB 421 TTCTGTGACTGTGGAGACACAGAGCATGGAAATCTGGCCCTTTTGTGTAAATCATGAA 480

QY 1197 CCTGGAAGCAGAGGTACTATAAAAGAGAAATTCACGCTGCTCGGTGGAATGGAAGAGGTAAAT 1256
DB 481 CCTGGAAGCAGAGGTACTATAAAAGAGAAATTCACGCTGCTCGGTGGAATGGAAGAGGTAAAT 540

QY 1257 GTCCAGCCAGGAAATATTTCTTCTAGTATAAAATATGCTGTAGAAATGACTATATGG 1316
DB 541 GTCCAGCCAGGAAATATTTCTTCTAGTATAAAATATGCTGTAGAAATGACTATATGG 600

QY 1317 GAAGAGGAAAGAACTCCCTCCCTGAACTCCAGATAAGGGAGAAATGAAAGATACTAT 1376
DB 601 GAAGAGGAAAGAACTCCCTCCCTGAACTCCAGATAAGGGAGAAATGAAAGATACTAT 660

QY 1377 TGTGTCTCTTTTCAATGATGAACACACATTTATATGACCAAGTCTATATACAGCCTACAAAGA 1436
DB 661 TGTGTCTCTTTTCAATGATGAACACACATTTATATGACCAAGTCTATATACAGCCTACAAAGA 720

QY 1437 GCTCTTGTGACTGTGAGCTCGCAGAGGCCAGTTGTCATACACATGCCCATTGACAAAGAGGT 1496
DB 721 GCTCTTGTGACTGTGAGCTCGCAGAGGCCAGTTGTCATACACATGCCCATTGACAAAGAGGT 780

QY 1497 CGTGGGCTGTTTAAAGCGGAGCTTATGCTGCTTCCAGGAGCAAAAGGAAAGATATAAG 1556
DB 781 CGTGGGCTGTTTAAAGCGGAGCTTATGCTGCTTCCAGGAGCAAAAGGAAAGATATAAG 840

QY 1557 AGTCATTCAGAAAATGCTCTCAACATCCACTTCATGTAGAGATATATACACTCAGAGATT 1616
DB 841 AGTCATTCAGAAAATGCTCTCAACATCCACTTCATGTAGAGATATATACACTCAGAGATT 900

QY 1617 ATGGCTCATCAGAAATTTGCTTTGGCTCTTGGCTCTTGGCTCTTGGCTCTTGGCTCTTGGCTCT 1676
DB 901 ATGGCTCATCAGAAATTTGCTTTGGCTCTTGGCTCTTGGCTCTTGGCTCTTGGCTCTTGGCTCT 960

QY 1677 TCAAGTGACTTTAGGCAGATCTTTTGGCAAGCATGCTTTAGAGAGAAACCTGACTCGGAG 1736
DB 961 TCAAGTGACTTTAGGCAGATCTTTTGGCAAGCATGCTTTAGAGAGAAACCTGACTCGGAG 1020

QY 1737 AATCCTCTCTCATAGCAGGTTAATGCTTTGGGATGCAAAAGCTTTTAAAGGTCCTCGT 1796
DB 1021 AATCCTCTCTCATAGCAGGTTAATGCTTTGGGATGCAAAAGCTTTTAAAGGTCCTCGT 1080

QY 1797 AAGATCCTTTCATGAATTTGCTTGGCTCTTGGCTCTTGGCTCTTGGCTCTTGGCTCTTGGCTCT 1856
DB 1081 AAGATCCTTTCATGAATTTGCTTGGCTCTTGGCTCTTGGCTCTTGGCTCTTGGCTCTTGGCTCT 1140

QY 1857 TTTGCTATGGAATTTGTGAAGTATTTATAAACAACTGCAAGAAAGATATATCAGTGATGAT 1916
DB 1141 TTTGCTATGGAATTTGTGAAGTATTTATAAACAACTGCAAGAAAGATATATCAGTGATGAT 1200

QY 1917 CATGACAGAGTATCTCTATAACTGCACTTTTCAGTTTCAGATGTTTACTGTTCTCTACTCTG 1976
DB 1201 CATGACAGAGTATCTCTATAACTGCACTTTTCAGTTTCAGATGTTTACTGTTCTCTACTCTG 1260

QY 1977 GCTCGACATCTTATGAGAGCAGAAATGTTATCTGTCATTTCTGTCATTTCTGTCATTTCTGTCAT 2036
DB 1261 GCTCGACATCTTATGAGAGCAGAAATGTTATCTGTCATTTCTGTCATTTCTGTCATTTCTGTCAT 1320

QY 2037 GTTTTACTCTGAGTACTTGGACAGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2096
DB 1321 GTTTTACTCTGAGTACTTGGACAGGAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1380

QY 2097 AAATTTGGAGAGATATATGCACTTAAAGTATATCTGATCAGCAAAACCC 2156
DB 1381 AAATTTGGAGAGATATATGCACTTAAAGTATATCTGATCAGCAAAACCC 1440

QY 2157 ACAATATGGACAGAAAGTTAAGATGCACTTCTGAGGTTTCTGAGGTTTCTGAGGTTTCTGAGG 2216
DB 1441 ACAATATGGACAGAAAGTTAAGATGCACTTCTGAGGTTTCTGAGGTTTCTGAGGTTTCTGAGG 1500

QY 2217 ATTCTTACTCTATGTCAGGAAATCGAAGAAATCGAAGACAGGTTGGGCAACACATTCGA 2276
DB 1501 ATTCTTACTCTATGTCAGGAAATCGAAGAAATCGAAGACAGGTTGGGCAACACATTCGA 1560

QY 2277 GTGGATCTGATTTGGAGGCTGCCATTTGCTATACAGATGCAATTAAGAAATTTTACTCTC 2336
DB 1561 GTGGATCTGATTTGGAGGCTGCCATTTGCTATACAGATGCAATTAAGAAATTTTACTCTC 1620

QY 2337 ATGTTTCCAGAGTGGTGTCTTGTGATGAGAACTTCTTACTTGTGGCTTATTAAGAAATGT 2396
DB 1621 ATGTTTCCAGAGTGGTGTCTTGTGATGAGAACTTCTTACTTGTGGCTTATTAAGAAATGT 1680

QY 2397 CACAAAGCTGTGAGGTGAGTACCAAGTTTCATATCTAGTAGCAAGACAGTAGTAGTACAA 2456
DB 1681 CACAAAGCTGTGAGGTGAGTACCAAGTTTCATATCTAGTAGCAAGACAGTAGTAGTACAA 1740

QY 2457 TCGTGTGACATAGTTTGGGAAACAAAGTCTTACAGAGTATCTGAGGATCTTGTGAAGCAT 2516

[illegible]

2821	TTGGGAAGTTCCAGCATGAATATACAAATGCTTTTGGAAAAACTCAAAGGAATTCCTCCAG	2888
3597	TTAAGAGGCGCAGAAGGACATGATAACGCTGGATACTTCAGATGTTTGACACAGTGAAGCGA	3656
2881	TTAAGAGGCGCAGAAGGACATGATAACGCTGGATACCTTCAGATGTTTGACACAGTGAAGCGA	2940
3657	TTAAGAGAAAAATCTTGTTTAATTGTAGCAACCAATCAGGATCGGAATCTATTAAAGAT	3716
2941	TTAAGAGAAAAATCTTGTTTAATTGTAGCAACCAATCAGGATCGGAATCTATTAAAGAT	3000
3717	GATGAGATTACTCATATAAGAAAAAGCAGAACGAAAAAGAAAAAGCTGAAGCTGCTAGG	3776
3001	GATGAGATTACTCATATAAGAAAAAGCAGAACGAAAAAGAAAAAGCTGAAGCTGCTAGG	3060
3777	CTACATCGCCAGAAAGATCATGGCTCAGATGTCTGCCCTTACAGAAAAAACTTCATTGAAACT	3836
3061	CTACATCGCCAGAAAGATCATGGCTCAGATGTCTGCCCTTACAGAAAAAACTTCATTGAAACT	3120
3837	CATAAACTCATGTPATGACAATAACATCAGAAAAATGCCCTGGGAAAGAGAAATTCATATTAGAG	3896
3121	CATAAACTCATGTPATGACAATAACATCAGAAAAATGCCCTGGGAAAGAGAAATTCATATTAGAG	3180
3897	GAGAGAGCACCCACGACGTCAGTGACTACTCTAGAAATTCCTTTGGGTCTCTAAACGGGGT	3956
3181	GAGAGAGCACCCACGACGTCAGTGACTACTCTAGAAATTCCTTTGGGTCTCTAAACGGGGT	3240
3957	CCATCTGTTACTGAAAAAGGAGGTGCTGACGTGCATCTCTTGCCAGAGAAACAGGAGGTG	4016
3241	CCATCTGTTACTGAAAAAGGAGGTGCTGACGTGCATCTCTTGCCAGAGAGAAACAGGAGGTG	3300
4017	AAAAATAGAAAAATATGCCATGTPATTATCGGCTGTGTCAGAAATCTACTGCTTTAAACC	4076
3301	AAAAATAGAAAAATATGCCATGTPATTATCGGCTGTGTCAGAAATCTACTGCTTTAAACC	3360
4077	CAGCACAGGGGAAAAACCATAGAACTCTCAGGAGAGCCCTAGACCCACTTTTTCATGGAT	4136
3361	CAGCACAGGGGAAAAACCATAGAACTCTCAGGAGAGCCCTAGACCCACTTTTTCATGGAT	3420
4137	CCAGACTTGGCATATCGAACTTATACAGGAAGCTGTGTCTCATGTAATGACGACAGTGTGC	4196
3421	CCAGACTTGGCATATCGAACTTATACAGGAAGCTGTGTCTCATGTAATGACGACAGTGTGC	3480
4197	TGGCAGAAGPATTTTGAAGCTGTACAGCTGAGCTCTCAGCAGCGCATTCATGTTGACCTT	4256
3481	TGGCAGAAGPATTTTGAAGCTGTACAGCTGAGCTCTCAGCAGCGCATTCATGTTGACCTT	3540
4257	TTTTGACTTGGAAAGTGGAGAAATCTTTGGCCCTCTTTCGCAATCTCTGTGCAATACGTG	4316
3541	TTTTGACTTGGAAAGTGGAGAAATCTTTGGCCCTCTTTCGCAATCTCTGTGCAATACGTG	3600
4317	ATCCCCATTATCTCTTTGCAACTCAAAAGATAAAACAGTGAGAAATCGAGATGCTCTTGCT	4376
3601	ATCCCCATTATCTCTTTGCAACTCAAAAGATAAAACAGTGAGAAATCGAGATGCTCTTGCT	3660
4377	CAACTTTTGACCTTGGCAGGTGGATACAGACTGTTCTTGGCCAGAAATATCAGGTTATAAT	4436
3661	CAACTTTTGACCTTGGCAGGTGGATACAGACTGTTCTTGGCCAGAAATATCAGGTTATAAT	3720
4437	ATAAGACATGCTAAAGGAGAAAAACCAATTCCTATTTTCTTTAAATCAAGGAATGGAGAT	4496
3721	ATAAGACATGCTAAAGGAGAAAAACCAATTCCTATTTTCTTTAAATCAAGGAATGGAGAT	3780
4497	TCTACTTTTGAGTTTCCATTCCTCATCTTGAGTTTTTGGGCTTTGAGTCTTCGATTAATAATTC	4556
3781	TCTACTTTTGAGTTTCCATTCCTCATCTTGAGTTTTTGGGCTTTGAGTCTTCGATTAATAATTC	3840
4557	AATAGCATCAAGGAAATTGGTTATTCTCTTTTGGCACAACAAATTTTATAGAAATGGAATGAA	4616
3841	AATAGCATCAAGGAAATTGGTTATTCTCTTTTGGCACAACAAATTTTATAGAAATGGAATGAA	3900
4617	GTGGCAGCTGATGAAAGGGATTCCTCGAGTCCCCATGCTGACCTGGAGACCTCGCGCTTTC	4676
3901	GTGGCAGCTGATGAAAGGGATTCCTCGAGTCCCCATGCTGACCTGGAGACCTCGCGCTTTC	3960

QY	4677	ACTATCCAGGCAATTGAAATCTATTGGGAGATGAAGGAAACCTCTCTGTTGGAGCACTT	4736
Db	3961	ACTATCCAGGCAATTGAAATCTATTGGGAGATGAAGGAAACCTCTCTGTTGGAGCACTT	4020
QY	4737	CAAAATAGCGACATAAATGGTCTGAAAGCATTAATGACAGTTGGCAGTTGCCAGAGGATT	4796
Db	4021	CAAAATAGCGACATAAATGGTCTGAAAGCATTAATGACAGTTGGCAGTTGCCAGAGGATT	4080
QY	4797	ACCTGCTCCTCAGGTCCTGATACAGAAACATCTGGTTGGTCTTCTATCAGTTGTTCTTCTCT	4856
Db	4081	ACCTGCTCCTCAGGTCCTGATACAGAAACATCTGGTTGGTCTTCTATCAGTTGTTCTTCTCT	4140
QY	4857	AACATAAATCAGAGATACACCATGCTTCTGCTCTATAGATCTGTTTCACTGTTTGGTG	4916
Db	4141	AACATAAATCAGAGATACACCATGCTTCTGCTCTATAGATCTGTTTCACTGTTTGGTG	4200
QY	4917	GGTGTGTTAGCATTCCTCCTCTGTTATTTGGGATGACCTGTTGATCTGCAGCCTTCT	4976
Db	4201	GGTGTGTTAGCATTCCTCCTCTGTTATTTGGGATGACCTGTTGATCTGCAGCCTTCT	4260
QY	4977	TCAGTTAGTCTTCTTAACACACCTTATCTCTTCCATTTGATCACCATGGCACACATG	5036
Db	4261	TCAGTTAGTCTTCTTAACACACCTTATCTCTTCCATTTGATCACCATGGCACACATG	4320
QY	5037	CTTCAGATACATCTTACAGTAGACACAGCCCTACCCCTTGCTCAGGTTCAAGAAGACAGT	5096
Db	4321	CTTCAGATACATCTTACAGTAGACACAGCCCTACCCCTTGCTCAGGTTCAAGAAGACAGT	4380
QY	5097	GAAGAGGCTCATTCGGCATCTCTTCTTCTGAGGAAATTTCTCAATATACAGTGGCTCC	5156
Db	4381	GAAGAGGCTCATTCGGCATCTCTTCTTCTGAGGAAATTTCTCAATATACAGTGGCTCC	4440
QY	5157	ATTGGGTGTGATATTCCTGGCTGGTATTTGCGGTCTCACTGAAGAAATGGCATCACCCCT	5216
Db	4441	ATTGGGTGTGATATTCCTGGCTGGTATTTGCGGTCTCACTGAAGAAATGGCATCACCCCT	4500
QY	5217	TATCTTCGCTGTGCTGATGTTGTTTCCATATTTACTTGGGGTAACTCCGCTCAGGAA	5276
Db	4501	TATCTTCGCTGTGCTGATGTTGTTTCCATATTTACTTGGGGTAACTCCGCTCAGGAA	4560
QY	5277	CTGCATACCAATCTGCAGAGGAGATACAGTGCACTCTGTAGCTATCTATCTTTACCT	5336
Db	4561	CTGCATACCAATCTGCAGAGGAGATACAGTGCACTCTGTAGCTATCTATCTTTACCT	4620
QY	5337	ACAAATTTGTTCTGCTCTCTTCAGCAATATTTGGGATCTGTAAGGCCCTTGCTCCAGAGG	5396
Db	4621	ACAAATTTGTTCTGCTCTCTTCAGCAATATTTGGGATCTGTAAGGCCCTTGCTCCAGAGG	4680
QY	5397	TGGTGTGCAGATCTCGCCTTACTAACTGTTTGAAGCAAAAAACACCGTGTGCAGGTAC	5456
Db	4681	TGGTGTGCAGATCTCGCCTTACTAACTGTTTGAAGCAAAAAACACCGTGTGCAGGTAC	4740
QY	5457	CCTAGAAAAAGAAATAGTTTGTAGAGTCTCTGATGATATAGTGCTGCTCTCTGATCAA	5516
Db	4741	CCTAGAAAAAGAAATAGTTTGTAGAGTCTCTGATGATATAGTGCTGCTCTCTGATCAA	4800
QY	5517	GCTTCTCATTTTCAGGTGCCCGGTCTGAGATGATGAGCAATCCCTGCTCTGCTG	5576
Db	4801	GCTTCTCATTTTCAGGTGCCCGGTCTGAGATGATGAGCAATCCCTGCTCTGCTG	4860
QY	5577	CTTTTCTGTGGGCTATATCTATGTTCTCAGAACATTTGCTGCAGGAAATTTGTAACGGG	5636
Db	4861	CTTTTCTGTGGGCTATATCTATGTTCTCAGAACATTTGCTGCAGGAAATTTGTAACGGG	4920
QY	5637	GAAGAGTTGGAGCTTGCAATTTTTCACGCATTTCTGCTGAGCCGAGTCTGCAATTTTC	5696
Db	4921	GAAGAGTTGGAGCTTGCAATTTTTCACGCATTTCTGCTGAGCCGAGTCTGCAATTTTC	4980
QY	5697	CTAAAAATCAGAGATGCGAGTGTCTGTTGAGGTAAAGCCAGAGGCTGTGCTCTAT	5756
Db	4981	CTAAAAATCAGAGATGCGAGTGTCTGTTGAGGTAAAGCCAGAGGCTGTGCTCTAT	5040

QY	5757	CCAGTCTTACTTGGATGAATATGGAGAAACAGACCCCTGGCTGAGAGGGGCAACCCC	5816
Db	5041	CCAGTCTTACTTGGATGAATATGGAGAAACAGACCCCTGGCTGAGAGGGGCAACCCC	5100
QY	5817	CTTCAATTTATCTCGTGAGCGGTATCGGAAGCTCCATTTGGTCTGGCAACAACACTGCAAT	5876
Db	5101	CTTCAATTTATCTCGTGAGCGGTATCGGAAGCTCCATTTGGTCTGGCAACAACACTGCAAT	5160
QY	5877	ATAGAAGAGATTGCTAGGAGCCAAAGAGACTAATCAGATGTTTATTTGGATTTCAACTGGCAG	5936
Db	5161	ATAGAAGAGATTGCTAGGAGCCAAAGAGACTAATCAGATGTTTATTTGGATTTCAACTGGCAG	5220
QY	5937	TTACTGTGAGCTCCAACTCTGAGGCTCAAGACAAATCAAAATGACGACAGTAGTAAAGGCTG	5996
Db	5221	TTACTGTGAGCTCCAACTCTGAGGCTCAAGACAAATCAAAATGACGACAGTAGTAAAGGCTG	5280
QY	5997	ATTCAAAATTTATGAAAACTTTCTGAGGCTGGGAAAGTATTTGGAGGCTCTTTTGTCTCCA	6056
Db	5281	ATTCAAAATTTATGAAAACTTTCTGAGGCTGGGAAAGTATTTGGAGGCTCTTTTGTCTCCA	5340
QY	6057	TGTCAGGTTCACTTACATCAATAAATAATTTCTTAATGGAGTATTTGCTTTCAATTAGCA	6116
Db	5341	TGTCAGGTTCACTTACATCAATAAATAATTTCTTAATGGAGTATTTGCTTTCAATTAGCA	5400
QY	6117	AACATATGCTTCACAGGAAAAAGACATAGATCAATCTGTTTATGCTAGTAGTATTTCC	6176
Db	5401	AACATATGCTTCACAGGAAAAAGACATAGATCAATCTGTTTATGCTAGTAGTATTTCC	5460
QY	6177	AGGAATTTATTTCCCTTCATATTTGCTCAATTTATTTTATTTTATTTTATTTTATTTT	6236
Db	5461	AGGAATTTATTTCCCTTCATATTTGCTCAATTTATTTTATTTTATTTTATTTTATTTT	5520
QY	6237	GAAGTCAGCTCAAAACAGTTGTAGACATTTTATGTTGTTGAACCTCTTCTGCAATTTGT	6296
Db	5521	GAAGTCAGCTCAAAACAGTTGTAGACATTTTATGTTGTTGAACCTCTTCTGCAATTTGT	5580
QY	6297	ATTTGGTGTGTTT 6308	
Db	5581	ATTTGGTGTGTTT 5592	
RESULT 4			
ID	AD116377	standard; DNA; 7648 BP.	
XX	AC	AD116377;	
XX	DT	22-APR-2004 (first entry)	
XX	DE	Human protein modification and maintenance molecule (PMM) gene #10.	
XX	KW	human; protein modification and maintenance molecule; PMM;	
XX	KW	gastrointestinal disorders; peptic ulcer; Crohn's disease;	
XX	KW	cardiovascular disorders; hypertension; congenital heart disease;	
XX	KW	autoimmune disease; inflammatory disease; AIDS; anaemia;	
XX	KW	developmental disorder; Cushing's syndrome; tubular acidosis;	
XX	KW	epithelial disorder; eczema; scabies; neurological disorder;	
XX	KW	Alzheimer's disease; multiple sclerosis; infection; cancer; gene; da.	
OS	XX	Homo sapiens.	
PN	XX	WO2003100016-A2.	
PD	XX	04-DEC-2003.	
XX	XX	22-MAY-2003; 2003WO-US016498.	
PF	XX	22-MAY-2002; 2002US-0383491P.	
PR	XX	24-JUN-2002; 2002US-0391378P.	
PR	XX	22-JUL-2002; 2002US-0397921P.	
XX	XX	(INCY-) INCYTE CORP.	
PA	XX		

PI Khare R, Bulloch SA, Swarnakar A, Elliott VS, Marquis JP;
PI Mason PM, Chawla NK, Ramkumar J, Kabile AE, Hafalia AJA, Lee SY;
PI Tran UK, Yue H, Becha SD, Griffin JA, Chang H, Jiang X, Jackson AA;
PI Richardson TW, Lal PG, Yao MG, Lu Y, Warren BA, Jin P, Wilson AD;
PI Gietzen KU;
XX WPI; 2004-035124/03.
DR P-PSDB; ADI16325.
XX
XX New protein modification and maintenance molecules, useful for diagnosing
PT or treating e.g. peptic ulcer, hypertension, rheumatic fever, AIDS,
PT Cushing's syndrome, Alzheimer's disease, multiple sclerosis, stroke or
PT cancers.
XX
XX Claim 5; SEQ ID NO 62; 419pp; English.
XX
XX The invention comprises the amino acid and coding sequences of human
CC protein modification and maintenance molecules (PMMM). The DNA and
CC protein sequences of the invention are useful for the diagnosis and
CC treatment of disorders associated with expression of PMMM, such as:
CC gastrointestinal disorders (e.g. peptic ulcer and Crohn's disease),
CC cardiovascular disorders (e.g. hypertension and congenital heart
CC disease), autoimmune or inflammatory disease (e.g. AIDS and anaemia),
CC developmental disorders (e.g. Cushing's syndrome and tubular acidosis),
CC epithelial disorders (e.g. eczema and scabies), neurological disorders
CC (e.g. Alzheimer's disease and multiple sclerosis), infections and cancer.
XX The present DNA sequence encodes a human PMMM protein of the invention.
XX
XX Sequence 7648 BP; 2355 A; 1407 C; 1606 G; 2280 T; 0 U; 0 Other;
XX
XX Query Match 85.4%; Score 5388; DB 12; Length 7648;
XX Best Local Similarity 98.2%; Pred. No. 0;
XX Matches 5498; Conservative 0; Mismatches 0; Indels 100; Gaps 1;
XX
XX 711 GCTGAGGTACTGACAGGATGAAATCAGCGGGAGTTACCCAGACCCCTCAGGCTG 770
XX 1 GCTGAGGTACTGACAGGATGAAATCAGCGGGAGTTACCCAGACCCCTCAGGCTG 60
XX
XX 771 GCATCTGGTGGATCAGCAAGTGAATTTATATCTGCTTTCTTGATCATTTGGCAAA 830
XX 61 GCATCTGGTGGATCAGCAAGTGAATTTATATCTGCTTTCTTGATCATTTGGCAAA 120
XX
XX 831 TTGGTGCAGAAATTTACTTTCTCAATGGACCCAGACTTGGAAAGCAGGAGGAAAGT 890
XX 121 TTGGTGCAGAAATTTACTTTCTGCAATGGACCCAGACTTGGAAAGCAGGAGGAAAGT 180
XX
XX 891 GTACAAATGTCAATATTCATCCACTGGAATGGTACTTATTTGGAGAGATCCAGATAT 950
XX 181 GTACAAATGTCAATATTCATCCACTGGAATGGTACTTATTTGGAGAGATCCAGATAT 240
XX
XX 951 TGCTTAGAGAAATTCAGACAGTGGAGCATTTGAGCTTTGGGAGGGTTTCAAAAGT 1010
XX 241 TGCTTAGAGAAATTCAGACAGTGGAGCATTTGAGCTTTGGGAGGGTTTCAAAAGT 300
XX
XX 1011 GGAGAGACAACTTATTTGAGGAGATTTGCAATTTGATCCAAATGTTACTCTGTATG 1070
XX 301 GGAGAGACAACTTATTTGAGGAGATTTGCAATTTGATCCAAATGTTACTCTGTATG 360
XX
XX 1071 GACTGCTTCCAGGACAGTGTTCATAAAAATCATCGTTTCAAGATGCATCTTCTACTGGA 1130
XX 361 GACTGCTTCCAGGACAGTGTTCATAAAAATCATCGTTTCAAGATGCATCTTCTACTGGA 420
XX
XX 1131 GGAGGGTTCTGTGACTGTGGAGACACAGAGGCATGGAAATCTGGCCCTTTTGTGTAAT 1190
XX 421 GGAGGGTTCTGTGACTGTGGAGACACAGAGGCATGGAAATCTGGCCCTTTTGTGTAAT 480
XX
XX 1191 CATGAACCTGGAAGACAGGTACTATTAAGAGAAATTCAGCTGTCCGTTGTAATGAAGAG 1250
XX 481 CATGAACCTGGAAGACAGGTACTATTAAGAGAAATTCAGCTGTCCGTTGTAATGAAGAG 540
XX
XX 1251 GTAATTTGTCAGCCAGGAAATATTTCTTCCAGTGATAAAAATATGTGCTAGAAATGACT 1310
XX 541 GTAATTTGTCAGCCAGGAAATATTTCTTCCAGTGATAAAAATATGTGCTAGAAATGACT 600

QY 1311 ATATGGAGAGGAAAAAGAACTGCTCTCTGAACTCCAGATAAGGAGAGAAAAATGAAGA 1370
DB 601 ATATGGAGAGGAAAAAGAACTGCTCTCTGAACTCCAGATAAGGAGAGAAAAATGAAGA 660
QY 1371 TACTATTGTGCTCTTTTCAATGATGAACACCAATTCATATGACCAAGTATATACAGCTA 1430
DB 661 TACTATTGTGCTCTTTTCAATGATGAACACCAATTCATATGACCAAGTATATACAGCTA 720
QY 1431 CAAAGAGCTTTGACTGTGAGCTCGAGAGCCGAGTTGCATACCACTGCCATTCGACAAA 1490
DB 721 CAAAGAGCTTTGACTGTGAGCTCGAGAGCCGAGTTGCATACCACTGCCATTCGACAAA 780
QY 1491 GAGGCTCGTGGGCTGTTAAAGCGGGAGCTTATCTCTGCTGCGAGGAGCAAGCAAGAGAT 1550
DB 781 GAGGCTCGTGGGCTGTTAAAGCGGGAGCTTATCTCTGCTGCGAGGAGCAAGCAAGAGAT 840
QY 1551 ATAAAGAGTCAATCAGAAAAATGCTCTCAACATCCACTTCATGTAAGTATATACACTCA 1610
DB 841 ATAAAGAGTCAATCAGAAAAATGCTCTCAACATCCACTTCATGTAAGTATATACACTCA 900
QY 1611 GAGATTATGGCTCATCAGAAATTTGCTTTGCGTCTTGCTTCTGATGAAACAAATATATG 1670
DB 901 GAGATTATGGCTCATCAGAAATTTGCTTTGCGTCTTGCTTCTGATGAAACAAATATATG 960
QY 1671 AGCTATTCAAGTACTTTAGGCAGATCTTTTGGCAAGCATGCTTTAGAGAGAAACCTGAC 1730
DB 961 AGCTATTCAAGTACTTTAGGCAGATCTTTTGGCAAGCATGCTTTAGAGAGAAACCTGAC 1020
QY 1731 TCGAGAGATCCCTGCTCTATAAGCAGGTTAATGCTTTGGGATGCAAGCTTTTATAAGGT 1790
DB 1021 TCGAGAGATCCCTGCTCTATAAGCAGGTTAATGCTTTGGGATGCAAGCTTTTATAAGGT 1080
QY 1791 GCCCGTAGATCCCTTCATGAAATTCATCTCAGCAGTTTCTTTTATGAGAGTGGATACAAA 1850
DB 1081 GCCCGTAGATCCCTTCATGAAATTCATCTCAGCAGTTTCTTTTATGAGAGTGGATACAAA 1140
QY 1851 AAACCTTTGCTATGGAATTTGCAAGTATTTATAAACAACCTGCGAGAAAGATATATCAGT 1910
DB 1141 AAACCTTTGCTATGGAATTTGCAAGTATTTATAAACAACCTGCGAGAAAGATATATCAGT 1200
QY 1911 GATGATCATGACAGAGTATCTTATAACCTGCACTTTTCAGTTCAGATGTTTACCTGTTCT 1970
DB 1201 GATGATCATGACAGAGTATCTTATAACCTGCACTTTTCAGTTCAGATGTTTACCTGTTCT 1260
QY 1971 ACTCTGGCTCGACATCTTTTGAAGCAGAAATGTTATCTCTGTCATTTACTGAAACTCTG 2030
DB 1261 ACTCTGGCTCGACATCTTTTGAAGCAGAAATGTTATCTCTGTCATTTACTGAAACTCTG 1320
QY 2031 CTAGAAGTTTCTGCTGAGTACTTGGACAGGAAACAATAAATTCACCTCCAGGGTTATAGC 2090
DB 1321 CTAGAAGTTTCTGCTGAGTACTTGGACAGGAAACAATAAATTCACCTCCAGGGTTATAGC 1380
QY 2091 CAGACAAATTTGGGAGAGTATATGCGATATATGTCGACTTAAAGTATATCTGATCAGC 2150
DB 1381 CAGACAAATTTGGGAGAGTATATGCGATATATGTCGACTTAAAGTATATCTGATCAGC 1440
QY 2151 AAACCCACATATGACAGAGAAATTAAGATGAGTTCCTTTGAGGTTTTCGATCTTTT 2210
DB 1441 AAACCCACATATGACAGAGAAATTAAGATGAGTTCCTTTGAGGTTTTCGATCTTTT 1500
QY 2211 TTGAAGATTTCTTACTGATGACAGAGAAATTAAGATGAGTTCCTTTGAGGTTTTCGATCTTTT 2270
DB 1501 TTGAAGATTTCTTACTGATGACAGAGAAATTAAGATGAGTTCCTTTGAGGTTTTCGATCTTTT 1560
QY 2271 ATTGAAGTGGATCTGATTTGGAGGCTGCCATTCATATACAGATGCAATTTGAAGATAT 2330
DB 1561 ATTGAAGTGGATCTGATTTGGAGGCTGCCATTCATATACAGATGCAATTTGAAGATAT 1620
QY 2331 TTACTCATGTTTCCAGAGTGGTGTGTCATGAGAACTCTTACTTTGCTGCTTATATA 2390
DB 1621 TTACTCATGTTTCCAGAGTGGTGTGTCATGAGAACTCTTACTTTGCTGCTTATATA 1680

2391 GAATGTCACAAAGCTGTGATGAGGTGCGAGTACCAGTTTCATATCTAGTAGCAAGCAGTA 2450
1681 GAATGTCACAAAGCTGTGATGAGGTGCGAGTACCAGTTTCATATCTAGTAGCAAGCAGTA 1740
2451 GTACAAATCGTGTGACATAGTTTGGAAACAAAGTCTTACAGAGTATCTGAGGATCTTGTGA 2510
1741 GTACAAATCGTGTGACATAGTTTGGAAACAAAGTCTTACAGAGTATCTGAGGATCTTGTGA 1800
2511 AGCATACATCTGCGCACTCTCTAGGACCTTGTGCTGCTCTTCAATGATGAGGCTG 2570
1801 AGCATACATCTGCGCACTCTCTAGGACCTTGTGCTGCTCTTCAATGATGAGGCTG 1860
2571 GGTGCTGTTTCAAGACTGCAATGATTTGCTGCTTGTGAGGACTTTCAAGTAGAGGTACTA 2630
1861 GGTGCTGTTTCAAGACTGCAATGATTTGCTGCTTGTGAGGACTTTCAAGTAGAGGTACTA 1920
2631 GTGGAATATCTTACGTTGTCTGCTGTTGTTGCTGCCAGGTGTTGCTGAGATGTTGGCGA 2690
1921 GTGGAATATCTTACGTTGTCTGCTGTTGTTGCTGCCAGGTGTTGCTGAGATGTTGGCGA 1980
2691 AGAAATGGAATGCTCTTTATAGCAGGTGTTTATTAACCAAGATGTTTAAAGTGCAGAGAA 2750
1981 AGAAATGGAATGCTCTTTATAGCAGGTGTTTATTAACCAAGATGTTTAAAGTGCAGAGAA 2040
2751 GAAATGATGATTAAGATATCATCATGCTTCAGATTTGCTGATGCTTTTAAATGATCCCAAT 2810
2041 GAAATGATGATTAAGATATCATCATGCTTCAGATTTGCTGATGCTTTTAAATGATCCCAAT 2100
2811 AAGTTCTTGTACTGTTACTTCAAGAGTATGAACTTGTGCGAGGCTTTTAAACAGACCATA 2870
2101 AAGTTCTTGTACTGTTACTTCAAGAGTATGAACTTGTGCGAGGCTTTTAAACAGACCATA 2160
2871 TCTACAAAGACCGAGATTGATTAACCAATATATATACACTAATAAGAAATGCTTCAG 2930
2161 TCTACAAAGACCGAGATTGATTAACCAATATATATACACTAATAAGAAATGCTTCAG 2220
2931 GTCTCTCATATATGTTGGTGTGAGGTTATGTACTGAGGTGGGAAATGTACCAAGAA 2990
2221 GTCTCTCATATATGTTGGTGTGAGGTTATGTACTGAGGTGGGAAATGTACCAAGAA 2280
2991 GAGGTCACAATGAGAGAAATCAATTCATCTTGTGATTTGAACCCATGCCACACAGTGCC 3050
2281 GAGGTCACAATGAGAGAAATCAATTCATCTTGTGATTTGAACCCATGCCACACAGTGCC 2340
3051 ATTGCGAAAAATTTACCTGAGAAATGAATTAATGAATGCTGCTTGAAGAAATGTCATAAC 3110
2341 ATTGCGAAAAATTTACCTGAGAAATGAATTAATGAATGCTGCTTGAAGAAATGTCATAAC 2400
3111 AAAGTGGCCACATTTAAGAAACCGAGTGTATCAGGCCATGAGGTTTATGAATCTAAAGAT 3170
2401 AAAGTGGCCACATTTAAGAAACCGAGTGTATCAGGCCATGAGGTTTATGAATCTAAAGAT 2460
3171 GAATCAGTGAAGACTTCAATATGTACTTTTATCAATCTCCAAACCCAGCATAGCAAG 3230
2461 GAATCAGTGAAGACTTCAATATGTACTTTTATCAATCTCCAAACCCAGCATAGCAAG 2520
3231 GCTGAAACATATGCAAGAAAGGAGAAACCAAGAAACCAAGATGAGCAATGCGGCCA 3290
2521 GCTGAAACATATGCAAGAAAGGAGAAACCAAGAAACCAAGATGAGCAATGCGGCCA 2580
3291 CCACACCTCTGAAATTCCTGCTGCTTTCAGCAAAAGTATTAACCTTCTCAACTGTGAT 3350
2581 CCACACCTCTGAAATTCCTGCTGCTTTCAGCAAAAGTATTAACCTTCTCAACTGTGAT 2640
3351 ATCATGATGATACATCTCAGGACCGTATTTTCAGGCGGCAATAGACACAGATTTCAACTTG 3410
2641 ATCATGATGATACATCTCAGGACCGTATTTTCAGGCGGCAATAGACACAGATTTCAACTTG 2700
3411 TGGACCGAAGGAGTGTCTCAAAATGCTTTTTCATATCTTGGCAATGGGTTTACTAGAGAG 3470
2701 TGGACCGAAGGAGTGTCTCAAAATG----- 2724
3471 AAGCAACAGCTTCAAAAGCTCTGAAAGAGAGTAACATTTTGTGACTTTTATCATTAAGGCT 3530

2725 ----- 2724
3531 TCAAGATTGGGAAGTTTCAAGCCATGAATATACAAATGCTTTTGGAAAACTCAAAGGAAT 3590
2725 ---GATTGGGAAGTTTCAAGCCATGAATATACAAATGCTTTTGGAAAACTCAAAGGAAT 2780
3591 CCCAGTTAGAAAGCCGAGAGGACATGATAAGCTGGATACATTCAGATGTTTGACACAGTG 3650
2781 CCCAGTTAGAAAGCCGAGAGGACATGATAAGCTGGATACATTCAGATGTTTGACACAGTG 2840
3651 AAGCGATTAAAGAAAAAATCTTGTAACTTGTAGCAACACATCAGGATCGGAATCTAAT 3710
2841 AAGCGATTAAAGAAAAAATCTTGTAACTTGTAGCAACACATCAGGATCGGAATCTAAT 2900
3711 AAGAAATGATGAGATTACTCATGATAAAGAAAAAGCAGAAACAAAAAGCTGAAGCT 3770
2901 AAGAAATGATGAGATTACTCATGATAAAGAAAAAGCAGAAACAAAAAGCTGAAGCT 2960
3771 GCTAGGCTACATCCGAGAGATCATGGCTCAGATGCTGCTTACAGAAAAAATCTTCAAT 3830
2961 GCTAGGCTACATCCGAGAGATCATGGCTCAGATGCTGCTTACAGAAAAAATCTTCAAT 3020
3831 GAAACTCATAACTCATGTATGACATACATCAGAAATGCTTGGAAAGAGATTCCTAAT 3890
3021 GAAACTCATAACTCATGTATGACATACATCAGAAATGCTTGGAAAGAGATTCCTAAT 3080
3891 ATGAGGAAGAGAGACCCCAAGCAGTCACTCTAGAAATTTGCTTTGGTCTCTAAA 3950
3081 ATGAGGAAGAGAGACCCCAAGCAGTCACTCTAGAAATTTGCTTTGGTCTCTAAA 3140
3951 CGGGGTCCATCTGTTACTGAAAAGAGGTGCTGACGTGATCTTTGCAAGAGAAACAG 4010
3141 CGGGGTCCATCTGTTACTGAAAAGAGGTGCTGACGTGATCTTTGCAAGAGAAACAG 3200
4011 GAGGTGAAAATAGAAAATAATGCCATGATATTCGGCTGTGTCAGAAATCTACTGCC 4070
3201 GAGGTGAAAATAGAAAATAATGCCATGATATTCGGCTGTGTCAGAAATCTACTGCC 3260
4071 TTAACCCAGCAGCAGGGGAAAACCCATAGAACTCTCAGGAGAAAGCCCTAGACCCACTTTTC 4130
3261 TTAACCCAGCAGCAGGGGAAAACCCATAGAACTCTCAGGAGAAAGCCCTAGACCCACTTTTC 3320
4131 ATGATCAGACTTGGCATATGGAATCTTATACAGAAAGCTGTGTGATGATGACAGCA 4190
3321 ATGATCAGACTTGGCATATGGAATCTTATACAGAAAGCTGTGTGATGATGACAGCA 3380
4191 GTGTGCTGGCAGAGTATTTGAAGCTGTACAGCTGAGCTCTCAGCAGGCAATTCATGTT 4250
3381 GTGTGCTGGCAGAGTATTTGAAGCTGTACAGCTGAGCTCTCAGCAGGCAATTCATGTT 3440
4251 GACCTTTTGAATTTGGAAGTGGAGAAATATCTTTGCCCTCTTTGCAATCTCTGTGCAAT 4310
3441 GACCTTTTGAATTTGGAAGTGGAGAAATATCTTTGCCCTCTTTGCAATCTCTGTGCAAT 3500
4311 ACTGTGATCCCAATTTATCTTTGCAACCTCAAAGATATAACAGTGAAGATGCAAGTGT 4370
3501 ACTGTGATCCCAATTTATCTTTGCAACCTCAAAGATATAACAGTGAAGATGCAAGTGT 3560
4371 CTGTGCTCAACTTTTGAACCTGGCAGGTGGATACAGACTGTTCTGGCCAGAAATATCAGGT 4430
3561 CTGTGCTCAACTTTTGAACCTGGCAGGTGGATACAGACTGTTCTGGCCAGAAATATCAGGT 3620
4431 TATAATATAGACATGCTTAAAGAGAAAAACCAATTTCTATTTTCTTTAAATCAAAGGAATG 4490
3621 TATAATATAGACATGCTTAAAGAGAAAAACCAATTTCTATTTTCTTTAAATCAAAGGAATG 3680
4491 GGAGATTTACTTTGGAGTTTCCATTTCCATCTGAGTTTGGCGTTGAGTCTTCAGATTA 4550
3681 GGAGATTTACTTTGGAGTTTCCATTTCCATCTGAGTTTGGCGTTGAGTCTTCAGATTA 3740
4551 TATTCAAATAGCATCAAGGAAATGGTTATTTCTTTTGGCACAACAAATTTATAGAAATGGA 4610

Db 3741 TATTCAATAGCATCAAGAAATGGTTATTCTCTTTGGCCACAACAATTTATAGAAATTGGA 3800
Qy 4611 TTGAAAGTGCACCTGATGAAAGGATCCTCGAGTCCCATGCTGACCTGGAGCACCTGC 4670
Db 3801 TTGAAAGTGCACCTGATGAAAGGATCCTCGAGTCCCATGCTGACCTGGAGCACCTGC 3860
Qy 4671 GCTTTCACTATTCAGGCAATTTGMAAATCTATTGGGAGATGAAGGAAACCTCTGTTTGG 4730
Db 3861 GCTTTCACTATTCAGGCAATTTGMAAATCTATTGGGAGATGAAGGAAACCTCTGTTTGG 3920
Qy 4731 GCATTTCAAAATAGGACGATTAATGGTCTGAAAGCATTAATGCAAGTTTGGCAGTTGCACAG 4790
Db 3921 GCATTTCAAAATAGGACGATTAATGGTCTGAAAGCATTAATGCAAGTTTGGCAGTTGCACAG 3980
Qy 4791 AGGATTACCTGCTCAGGTCCTGATACAGAAACATCTGGTTGGTCTCTCTATCAGTTGTT 4850
Db 3981 AGGATTACCTGCTCAGGTCCTGATACAGAAACATCTGGTTGGTCTCTCTATCAGTTGTT 4040
Qy 4851 CTTCCTAACATAAATCAGAAGATACACATGCCCTTCTGTCTATAGATCTGTTTCATGTT 4910
Db 4041 CTTCCTAACATAAATCAGAAGATACACATGCCCTTCTGTCTATAGATCTGTTTCATGTT 4100
Qy 4911 TTGGTGGGTGCTGTGTAGCATTCCTCATCTTGTATTGGGATGACCCCTGTTGATCTGCAG 4970
Db 4101 TTGGTGGGTGCTGTGTAGCATTCCTCATCTTGTATTGGGATGACCCCTGTTGATCTGCAG 4160
Qy 4971 CTTCTTCAGTTAGTTCTTCTTAACACACTTATCTCTTCCATTTGATCACCATTGGCA 5030
Db 4161 CTTCTTCAGTTAGTTCTTCTTAACACACTTATCTCTTCCATTTGATCACCATTGGCA 4220
Qy 5031 CACATGCTTCAGATACTACTTACAGTACAGCCCTTACCCCTTGCTCAGGTTCAAGAA 5090
Db 4221 CACATGCTTCAGATACTACTTACAGTACAGCCCTTACCCCTTGCTCAGGTTCAAGAA 4280
Qy 5091 GACAGTGAAGAGGCTCATTCGGCATCTCTTTCTTTGAGAAATTTCTCAATATACAGT 5150
Db 4281 GACAGTGAAGAGGCTCATTCGGCATCTCTTTCTTTGAGAAATTTCTCAATATACAGT 4340
Qy 5151 GGCTCCATGGGTGATGATTCCTGGCTGGTATTGTTGGGGTCTCACTGAAGAAATGGCATC 5210
Db 4341 GGCTCCATGGGTGATGATTCCTGGCTGGTATTGTTGGGGTCTCACTGAAGAAATGGCATC 4400
Qy 5211 ACCCTTATCTTCGCTGCTGCATGTTGTTTCCACTATTACTTTGGGGTAACCTCCGCT 5270
Db 4401 ACCCTTATCTTCGCTGCTGCATGTTGTTTCCACTATTACTTTGGGGTAACCTCCGCT 4460
Qy 5271 GAGGAACCTGCATACCAATTTCTGCAAGAGGAGATACAGTGCACCTCTGTAGCTATCTATCT 5330
Db 4461 GAGGAACCTGCATACCAATTTCTGCAAGAGGAGATACAGTGCACCTCTGTAGCTATCTATCT 4520
Qy 5331 TTACCTACAAATTTGTTCTGCTCTTCCAGGAATATTGGGATACATGTAAGGCCCTTGCTC 5390
Db 4521 TTACCTACAAATTTGTTCTGCTCTTCCAGGAATATTGGGATACATGTAAGGCCCTTGCTC 4580
Qy 5391 CAGAGTGGTGGCAGATCCTGCCCTTACTAACTGTTTGAAGCAAAAAACACCGTGTCT 5450
Db 4581 CAGAGTGGTGGCAGATCCTGCCCTTACTAACTGTTTGAAGCAAAAAACACCGTGTCT 4640
Qy 5451 AGGTACCTTAGAAAAGAAATAGTTTGTATAGAGCTTCTCTGATGACTATAGCTGCCCTCTG 5510
Db 4641 AGGTACCTTAGAAAAGAAATAGTTTGTATAGAGCTTCTCTGATGACTATAGCTGCCCTCTG 4700
Qy 5511 AATCAAGCTTCTCATTTCAAGTGGCCACCGTCTGCAGATGATGACGCAAAAGCATCTGCTC 5570
Db 4701 AATCAAGCTTCTCATTTCAAGTGGCCACCGTCTGCAGATGATGACGCAAAAGCATCTGCTC 4760
Qy 5571 CTCTGCCCTTTCTGTTGGGCTATATCTATGTTCTCAGAACATTTGCTGCCAGGAATTTGT 5630
Db 4761 CTCTGCCCTTTCTGTTGGGCTATATCTATGTTCTCAGAACATTTGCTGCCAGGAATTTGT 4820
Qy 5631 AACGGGGAAGAGTTGGAGCTTGGCAATTTTTTTCACGCACTTCTCACTGTGGAGCCGAGTCTGC 5690
Db 4821 AACGGGGAAGAGTTGGAGCTTGGCAATTTTTTTCACGCACTTCTCACTGTGGAGCCGAGTCTGC 4880

Qy 5691 ATTTTCTTAAAAATCAGAGAAATGCCAGTGGTCTGTTGTAAGGTAAGCCAGAGGCTGT 5750
Db 4881 ATTTTCTTAAAAATCAGAGAAATGCCAGTGGTCTGTTGTAAGGTAAGCCAGAGGCTGT 4940
Qy 5751 GCCTATCCAGCTCCTTACTTGGATGAATATGGAGAAAACAGACCTCGCCTGGAAGGGGC 5810
Db 4941 GCCTATCCAGCTCCTTACTTGGATGAATATGGAGAAAACAGACCTCGCCTGGAAGGGGC 5000
Qy 5811 AACCCCTTCAATTTATCTCGTAGCGGTATCGAGAGCTCCATTTGGTCTGGCAACAACAC 5870
Db 5001 AACCCCTTCAATTTATCTCGTAGCGGTATCGAGAGCTCCATTTGGTCTGGCAACAACAC 5060
Qy 5871 TGCATTTAGAAGAGATTGCTTAGGAGCAAGAGACTAATCAGATGTTATTGGATTCAAC 5930
Db 5061 TGCATTTAGAAGAGATTGCTTAGGAGCAAGAGACTAATCAGATGTTATTGGATTCAAC 5120
Qy 5931 TGGCAGTTACTGTGAGCTCCAACTCTGCCTCAAGCAATCAAAATGACGACAGTAGTAA 5990
Db 5121 TGGCAGTTACTGTGAGCTCCAACTCTGCCTCAAGCAATCAAAATGACGACAGTAGTAA 5180
Qy 5991 AGGCTGATTTCAAAATTAAGGAAAACCTTTCTGAGGGCTGGGAAAGTATTGGAGGCTCTTT 6050
Db 5181 AGGCTGATTTCAAAATTAAGGAAAACCTTTCTGAGGGCTGGGAAAGTATTGGAGGCTCTTT 5240
Qy 6051 GCTCCATGTCAGGTTCACTTTACATCAATAAATAATTTCTTAATGGAGTATTGCTTTCAA 6110
Db 5241 GCTCCATGTCAGGTTCACTTTACATCAATAAATAATTTCTTAATGGAGTATTGCTTTCAA 5300
Qy 6111 TTAGCAAAACATATGCTTTACAGGAAAAAAGACATAGATCAATCTGTTTATGCTAGT 6170
Db 5301 TTAGCAAAACATATGCTTTACAGGAAAAAAGACATAGATCAATCTGTTTATGCTAGT 5360
Qy 6171 ATTTCCAGGAATTTATTTCCCTTCAATAATTTGCTCATTTCAATTTATTCATCCACTTG 6230
Db 5361 ATTTCCAGGAATTTATTTCCCTTCAATAATTTGCTCATTTCAATTTATTCATCCACTTG 5420
Qy 6231 GTAGATCAAGTTCAGTCAAAACAGTTGTAGACATTTTATGTTGTTAACTCTTCTGCAA 6290
Db 5421 GTAGATCAAGTTCAGTCAAAACAGTTGTAGACATTTTATGTTGTTAACTCTTCTGCAA 5480
Qy 6291 TTTTGTATTTGGTGTCTTT 6308
Db 5481 TTTTGTATTTGGTGTCTTT 5498

RESULT 5

ADJ95458
ID ADJ95458 standard; cDNA; 5205 BP.

XX AC ADJ95458;

XX XX 03-JUN-2004 (first entry)

XX XX Human cDNA encoding Ubiquitin ligase E3alpha I, synthetic variant.

XX XX Human; ss; gene; ubiquitin ligase; E3alpha I;

KW ubiquitin-proteasome pathway; gene therapy; vaccine; muscular atrophy;

KW cachexia; catabolic disorders; cancer cachexia; renal cachexia;

KW inflammatory cachexia; muscle wasting disorder; metabolic acidosis;

KW uremia; burn; hyperthyroidism; Cushing's syndrome; fasting;

KW denervation atrophy; diabetes mellitus; sepsis; AIDS wasting syndrome;

KW SNP; single nucleotide polymorphism.

XX XX Homo sapiens.

OS Synthetic.

XX XX Location/Qualifiers

PH 1..5205

FT /*tag= a

FT /product= "Amino acids 6-1734 of SEQ ID 19"

FT /transl_except= (pos:658..678,aa:REKNERYCYVLFNDEHHSHYDHV)

FT /transl_except= (pos:4933..4998,aa:CK)

```
FT variation replace(4657,T)
FT /*tag= b
XX /standard_name= "single nucleotide polymorphism"
PN US6706505-B1.
XX
XX 16-MAR-2004.
XX
XX 28-NOV-2000; 2000US-007241126.
XX
XX 08-MAR-2000; 2000US-0187911P.
XX
XX (AMGE-) AMGEN INC.
XX
XX Han H, Kwak K;
XX
XX WPI; 2004-236723/22.
XX
XX P-PSDB; ADJ95459.
XX
XX New nucleic acid molecule, useful for preparing a composition for
XX diagnosing, treating or preventing diseases associated with human
XX E3approximately polypeptide, e.g., muscle atrophy.
XX
XX Claim 19; SEQ ID NO 18; 104pp; English.
XX
XX The invention relates to a new isolated nucleic acid molecule appearing
XX as ADJ95441 (or its complement) encoding a ubiquitin ligase E3alpha I
XX protein appearing as ADJ95442. Also included are a vector comprising the
XX nucleic acid, a host cell comprising the vector, a process of producing a
XX E3alpha I ubiquitin ligase polypeptide, a composition comprising the
XX nucleic acid molecule, a reagent comprising a detectably labelled
XX nucleotide, and a method for determining the presence of a human E3alpha
XX I ubiquitin ligase nucleic acid in a biological sample. The nucleic acid
XX molecule is useful for preparing a composition for diagnosing, treating
XX or preventing diseases associated with human E3alpha I polypeptide, e.g.
XX muscle atrophy, cachexia, catabolic disorders, cancer cachexia, renal
XX cachexia, inflammatory cachexia, muscle wasting disorders associated with
XX metabolic acidosis, uremia, burns, hyperthyroidism, Cushing's syndrome,
XX fasting, denervation atrophy, diabetes mellitus, sepsis and AIDS wasting
XX syndrome. The present sequence encodes a human E3alpha I variant. NOTE:
XX This cDNA was assembled from previously isolated cDNA fragments and is
XX stated to encode ADJ95459, but appears to be incomplete as it only
XX encodes from amino acid 6 and has translation exceptions.
XX
XX Sequence 5205 BP; 1553 A; 1017 C; 1171 G; 1450 T; 0 U; 14 Other;
XX
XX Query Match 81.4%; Score 5136; DB 12; Length 5205;
XX Best Local Similarity 98.8%; Pred No. 0;
XX Matches 5189; Conservative 10; Mismatches 6; Indels 45; Gaps 1;
XX
XX 696 ATGCGGACGAGGAGGCTGGAGGTACTGAGAGGATGGAATCAGCGCGAGTTACCCAG 755
XX 1 ATGCGGACGAGGAGGCTGGAGGTACTGAGAGGATGGAATCAGCGCGAGTTACCCAG 60
XX
XX 756 ACCCTCAGCGTCTGGCATCTTGGTGGATCAGCAAGTTGATTTTATCTGCTTTCTTG 815
XX 61 ACCCTCAGCGTCTGGCATCTTGGTGGATCAGCAAGTTGATTTTATCTGCTTTCTTG 120
XX
XX 816 CATCATTTGGCACAATTTGGTCCAGAAATTTACTTTGCTGAAATGGACCCAGACTTGGAA 875
XX 121 CATCATTTGGCACAATTTGGTCCAGAAATTTACTTTGCTGAAATGGACCCAGACTTGGAA 180
XX
XX 876 AAGCAGGAGGAAAGTGTAATGTCAATATTTCACTCCACTGGAATGGTACTTATTTGGA 935
XX 181 AAGCAGGAGGAAAGTGTAATGTCAATATTTCACTCCACTGGAATGGTACTTATTTGGA 240
XX
XX 936 GAAGATCCAGATATTTGCTTTAGAGAAATTTAGAGCAAGTGGAGCATTTTTCAGTTTGGG 995
XX 241 GAAGATCCAGATATTTGCTTTAGAGAAATTTAGAGCAAGTGGAGCATTTTTCAGTTTGGG 300
XX
XX 996 AGGGTTTTCAAAAGTGGAGGACACCTATTCTTTCAGGGATTGCAATTTGATCCACACA 1055
XX 301 AGGGTTTTCAAAAGTGGAGGACACCTATTCTTTCAGGGATTGCAATTTGATCCACACA 360
```

```
QY 1056 TGTGTACTCTGTATGGACTGCTTCCAGGACAGTGTTCATATAAAATCATCGTTTACAAGATG 1115
DB 361 TGTGTACTCTGTATGGACTGCTTCCAGGACAGTGTTCATATAAAATCATCGTTTACAAGATG 420
QY 1116 CATACTTCTTACTTGGAGGAGGGTTCTGTCACTGTGTGAGACACAGAGGCAATGGAATACTGGC 1175
DB 421 CATACTTCTTACTTGGAGGAGGGTTCTGTCACTGTGTGAGACACAGAGGCAATGGAATACTGGC 480
QY 1176 CCTTTTGTGTAAATCATGAACCTTGGAAAGAGCAGGTACTATATAAAGAGAAATTCACGCTGT 1235
DB 481 CCTTTTGTGTAAATCATGAACCTTGGAAAGAGCAGGTACTATATAAAGAGAAATTCACGCTGT 540
QY 1236 CGTTGAATGAAGAGGTAAATTTGTCCAAGCCAGGAAAATATTTCTTCTAGTGATATAAATAT 1295
DB 541 CGTTGAATGAAGAGGTAAATTTGTCCAAGCCAGGAAAATATTTCTTCTAGTGATATAAATAT 600
QY 1296 GTCTGTAGAAATGCACTATATGGAAGAGGAAAGAACTGCTTCCCTGAACTCCAGATAGG 1355
DB 601 GTCTGTAGAAATGCACTATATGGAAGAGGAAAGAACTGCTTCCCTGAACTCCAGATAGG 660
QY 1356 GAGAAAAATGAAGATACCTATTGTGTCTCTTTTCAATGATGAACACCATTCATATACACAC 1415
DB 661 KNR-----YVCVNDHSHVDH 675
QY 1416 GTCATATACAGCCTACAAAGAGCTTTGACTGTGTGAGCTCGCAGAGGCCAGTTGCATACC 1475
DB 676 GTCATATACAGCCTACAAAGAGCTTTGACTGTGTGAGCTCGCAGAGGCCAGTTGCATACC 735
QY 1476 ACTGCCATTGACAAAGAGGGTCTCGGGCTGTGTAAGCGGGAGCTTATGCTGCTTGCCAG 1535
DB 736 ACTGCCATTGACAAAGAGGGTCTCGGGCTGTGTAAGCGGGAGCTTATGCTGCTTGCCAG 795
QY 1536 GAAGCAAAAGGAGATATAAAGAGTCATTCAGAAAAATGTCTCTCAACATCCACTTCATGTA 1595
DB 796 GAAGCAAAAGGAGATATAAAGAGTCATTCAGAAAAATGTCTCTCAACATCCACTTCATGTA 855
QY 1596 GAAGTATTACACTCAGAGATTTATGGCTCATCAGAAAAATTTGCTTTCGCTTGGTTCTGG 1655
DB 856 GAAGTATTACACTCAGAGATTTATGGCTCATCAGAAAAATTTGCTTTCGCTTGGTTCTGG 915
QY 1656 ATGAACAAAAATATGAGCTATTCAAGTGACTTTTAGGCGAGATCTTTTGGCCAGCATGCCCT 1715
DB 916 ATGAACAAAAATATGAGCTATTCAAGTGACTTTTAGGCGAGATCTTTTGGCCAGCATGCCCT 975
QY 1716 AGAGAGAACCTGACTCGGAGAACTCCCTGTCTCTAAGCAGGTTAAATGCTTTGGGATGCA 1775
DB 976 AGAGAGAACCTGACTCGGAGAACTCCCTGTCTCTAAGCAGGTTAAATGCTTTGGGATGCA 1035
QY 1776 AAGCTTTATAAAGGTGCCGTAAGATCCTTATGAATTTGATCTTCAGCAGTTTTTTTATG 1835
DB 1036 AAGCTTTATAAAGGTGCCGTAAGATCCTTATGAATTTGATCTTCAGCAGTTTTTTTATG 1095
QY 1836 GAGATGGATACAAAAAATCTTTTGTCTATGAAATTTTGTGAAGTATATAAACACTGCGAG 1895
DB 1096 GAGATGGATACAAAAAATCTTTTGTCTATGAAATTTTGTGAAGTATATAAACACTGCGAG 1155
QY 1896 AAAGAAATATATCAGTGTGATCATGACAGAAAGTATCTCTATAAATGCACTTTTCACTTCA 1955
DB 1156 AAAGAAATATATCAGTGTGATCATGACAGAAAGTATCTCTATAAATGCACTTTTCACTTCA 1215
QY 1956 ATGTTTACTGTTCTTACTCTGGCTCGACATCTTATTTGAAGCAGAAATGTTTATCTCTGTC 2015
DB 1216 ATGTTTACTGTTCTTACTCTGGCTCGACATCTTATTTGAAGCAGAAATGTTTATCTCTGTC 1275
QY 2016 ATTTACTGAATCTGCTAGAAATTTTACTGAGTACTTTGGACAGGAAATAAATTTCAAC 2075
DB 1276 ATTTACTGAATCTGCTAGAAATTTTACTGAGTACTTTGGACAGGAAATAAATTTCAAC 1335
QY 2076 TTCCAGGGTTATAGCAGGACAAAATTTGGGAAGAGTATATGCAAGTAAATATGTGACCTAAG 2135
DB 1336 TTCCAGGGTTATAGCAGGACAAAATTTGGGAAGAGTATATGCAAGTAAATATGTGACCTAAG 1395
```

QY	2136	TATATCTGATCAGCAAAACCCACAAATATGACAGAAAGATTAAAGATGACGCTTCTTTGAA	2195
DB	1396	TATATCTGATCAGCAAAACCCACAAATATGACAGAAAGATTAAAGATGACGCTTCTTTGAA	1455
QY	2196	GGTTTTGCGATCTTTTTGGAAGATTCTTACCTGTATGACGGAATGGAAGAAATCCGAAGA	2255
DB	1456	GGTTTTGCGATCTTTTTGGAAGATTCTTACCTGTATGACGGAATGGAAGAAATCCGAAGA	1515
QY	2256	CAGTTGGGCAACACATTTGAAGTGAATCCTGATTTGGGAGGCTGCCATGCTATACAGATG	2315
DB	1516	CAGTTGGGCAACACATTTGAAGTGAATCCTGATTTGGGAGGCTGCCATTTGCTATACAGATG	1575
QY	2316	CAATTTGAAGATAATTTTACTCATGTTCCAGAGTGTGTGCTGTGTGATGAAGAACTCTTAA	2375
DB	1576	CAATTTGAAGATAATTTTACTCATGTTCCAGAGTGTGTGCTGTGTGATGAAGAACTCTTAA	1635
QY	2376	CTTGTGGCTTATAAAGAAATGTACAAAGCTGTGTGATGAGGTGACGTACAGTTTTCATATCT	2435
DB	1636	CTTGTGGCTTATAAAGAAATGTACAAAGCTGTGTGATGAGGTGACGTACAGTTTTCATATCT	1695
QY	2436	AGTAGCAAGACAGTAGTACAAATCGTGTGGAATAGTTTGGAAACAAAGTCTTACAGAGTA	2495
DB	1696	AGTAGCAAGACAGTAGTACAAATCGTGTGGAATAGTTTGGAAACAAAGTCTTACAGAGTA	1755
QY	2496	TCTGAGGATCTTGTAAAGCATACATCTGCCACTCTCTAGGACCTTTGCTGGTCTTCATGTA	2555
DB	1756	TCTGAGGATCTTGTAAAGCATACATCTGCCACTCTCTAGGACCTTTGCTGGTCTTCATGTA	1815
QY	2556	CGTTTAAAGCAGGCTGGGTGCTGTTTCAAGACTGATGAATTTGTGTCTTTTGGAGACTTTT	2615
DB	1816	CGTTTAAAGCAGGCTGGGTGCTGTTTCAAGACTGATGAATTTGTGTCTTTTGGAGACTTTT	1875
QY	2616	CAAGTAGAGGTACTAGTGGGAATATCTTTTACGTTGCTGGTGTGGTGGCCAGGTTGTT	2675
DB	1876	CAAGTAGAGGTACTAGTGGGAATATCTTTTACGTTGCTGGTGTGGTGGCCAGGTTGTT	1935
QY	2676	GCTGAGATGTGGCAAGAAATGGACTGCTCTTATAGCCAGGTGTTTATTTACCAAGAT	2735
DB	1936	GCTGAGATGTGGCAAGAAATGGACTGCTCTTATAGCCAGGTGTTTATTTACCAAGAT	1995
QY	2736	GTTAAGTGCAGAGAAATGTATGATAAGATATCATGCTTTCAGATTTGGTGCACT	2795
DB	1996	GTTAAGTGCAGAGAAATGTATGATAAGATATCATGCTTTCAGATTTGGTGCACT	2055
QY	2796	TTAATGGAATCCCAATAGTCTTGTGTTACTGTTACTTTCAGAGGTATGAATTTGCCAGGCT	2855
DB	2056	TTAATGGAATCCCAATAGTCTTGTGTTACTGTTACTTTCAGAGGTATGAATTTGCCAGGCT	2115
QY	2856	TTTAAACAGACCATATCTACAAAGACCAAGCATTTGATTAAACAAATATAATACACTAATA	2915
DB	2116	TTTAAACAGACCATATCTACAAAGACCAAGCATTTGATTAAACAAATATAATACACTAATA	2175
QY	2916	GAAAGAAATGCTTCAGGCTCCTCATCTATATTTGGGTGAGCGTTATGTACCTGGAGTGGGA	2975
DB	2176	GAAAGAAATGCTTCAGGCTCCTCATCTATATTTGGGTGAGCGTTATGTACCTGGAGTGGGA	2235
QY	2976	AATGTGACAAAGAGAGGTCAAAATGAGAGAAATCATTTCACTGCTTTGCAATGGAACCC	3035
DB	2236	AATGTGACAAAGAGAGGTCAAAATGAGAGAAATCATTTCACTGCTTTGCAATGGAACCC	2295
QY	3036	ATGCCACACAGTCCATTTGCCAAAATTTTACTGAGAAATGAAATAATGAAACTGGCTTAA	3095
DB	2296	ATGCCACACAGTCCATTTGCCAAAATTTTACTGAGAAATGAAATAATGAAACTGGCTTAA	2355
QY	3096	GAGAAATGTCAATAAACAAGTGGCCACATTTTAAAGAACCAAGGTGTATCAGGCCATGGAGTT	3155
DB	2356	GAGAAATGTCAATAAACAAGTGGCCACATTTTAAAGAACCAAGGTGTATCAGGCCATGGAGTT	2415
QY	3156	TATGAACTAAAGATGAATCACTGAAAGACTTCAATATGTACTTTTATCATTTACTCCAAA	3215
DB	2416	TATGAACTAAAGATGAATCACTGAAAGACTTCAATATGTACTTTTATCATTTACTCCAAA	2475
QY	3216	ACCCAGCATAGCAAGGCTGAAATATATGACAGAAAGAGAGAAACAAAGAAACAAAGAT	3275

DB	2476	ACCAGCATAGCAAGGCTGAACATATGACAGAAAGAGAGAAACAAAGAAACAAAGAT	2535
QY	3276	GAAGCATTTGGCGCCACACACCTCTCGTAATTTCTGCCCTCTCTTTCAGCAAGTGAATTAAC	3335
DB	2536	GAAGCATTTGGCGCCACACACCTCTCGTAATTTCTGCCCTCTCTTTCAGCAAGTGAATTAAC	2595
QY	3336	CTTCTCAACTGTGATATCATGATGTATCATTTCTCAGGACCGTATTTGAGCGGGCAATAGAC	3395
DB	2596	CTTCTCAACTGTGATATCATGATGTATCATTTCTCAGGACCGTATTTGAGCGGGCAATAGAC	2655
QY	3396	ACAGATTTCTAACTTTGTGGACCGGAAGGGATGCTCCAAATGCTTTTTCATATTTCTGGCAATG	3455
DB	2656	ACAGATTTCTAACTTTGTGGACCGGAAGGGATGCTCCAAATGCTTTTTCATATTTCTGGCAATG	2715
QY	3456	GGTTTACTAGAGAGAGCAACAGCTTCAAAAAGCTCTGAGAGAGAGTAAACATTTGAC	3515
DB	2716	GGTTTACTAGAGAGAGCAACAGCTTCAAAAAGCTCTGAGAGAGAGTAAACATTTGAC	2775
QY	3516	TTTTATCATTAAGGCTTCAAGATTGGGAAGTTTCAAGCCATGAATATACAAATGCTTTTGGAA	3575
DB	2776	TTTTATCATTAAGGCTTCAAGATTGGGAAGTTTCAAGCCATGAATATACAAATGCTTTTGGAA	2835
QY	3576	AAACTCAAAAGGAAATTTCCCGAGTTTAGAAGGCCAGAGCAATGATAAAGTGGATACTTCAG	3635
DB	2836	AAACTCAAAAGGAAATTTCCCGAGTTTAGAAGGCCAGAGCAATGATAAAGTGGATACTTCAG	2895
QY	3636	ATGTTTGCACACAGTGAAGCGATTAAGAGAAATTTCTTTTAAATTTGATAGCAACCATCA	3695
DB	2896	ATGTTTGCACACAGTGAAGCGATTAAGAGAAATTTCTTTTAAATTTGATAGCAACCATCA	2955
QY	3696	GGATCGGAATCTATTAAGAATGATGAGATTACTCATGATAAGAAAGAGAGAGAGAGAGAG	3755
DB	2956	GGATCGGAATCTATTAAGAATGATGAGATTACTCATGATAAGAAAGAGAGAGAGAGAGAG	3015
QY	3756	AGAAAAGCTGAAGCTGTAGGCTACATCGCCAGAGAGATCATGGCTCAGATGTCTGCCCTTA	3815
DB	3016	AGAAAAGCTGAAGCTGTAGGCTACATCGCCAGAGAGATCATGGCTCAGATGTCTGCCCTTA	3075
QY	3816	CAGAAAACTTCAATGAACTCATATAAATCATGATCATGATGATGATGATGATGATGATGATG	3875
DB	3076	CAGAAAACTTCAATGAACTCATATAAATCATGATCATGATGATGATGATGATGATGATGATG	3135
QY	3876	AAAGAGATTTCCATTTATGGAG	3935
DB	3136	AAAGAGATTTCCATTTATGGAG	3195
QY	3936	GCTTTGGGTCTTAAACGGGGTCCATCTGTTTCTGAAAAGGAGGTGCTGACGTGATCCTT	3995
DB	3196	GCTTTGGGTCTTAAACGGGGTCCATCTGTTTCTGAAAAGGAGGTGCTGACGTGATCCTT	3255
QY	3996	TGCCAAG	4055
DB	3256	TGCCAAG	3315
QY	4056	CAGAAATCTACTGCTTAAACCCAGACAGAGGGGAAACCCATAGAACTCTCAGAGAGAGAGCC	4115
DB	3316	CAGAAATCTACTGCTTAAACCCAGACAGAGGGGAAACCCATAGAACTCTCAGAGAGAGAGCC	3375
QY	4116	CTAGACCCACTTTTTCATGGATCCAGACTTGGCATATGGAACTTATACAGGAAGCTGTGGT	4175
DB	3376	CTAGACCCACTTTTTCATGGATCCAGACTTGGCATATGGAACTTATACAGGAAGCTGTGGT	3435
QY	4176	CATGTAATGACGAGGTGTGCTGCGAGAGATTTTGAAGCTGTACAGCTGAGCTCTTCAG	4235
DB	3436	CATGTAATGACGAGGTGTGCTGCGAGAGATTTTGAAGCTGTACAGCTGAGCTCTTCAG	3495
QY	4236	CAGCGCATTCATGTTGACCTTTTGAATTTGAGAGTGGAGAAATCTTTTGGCCCTCTTTC	4295
DB	3496	CAGCGCATTCATGTTGACCTTTTGAATTTGAGAGTGGAGAAATCTTTTGGCCCTCTTTC	3555
QY	4296	AAATCTCTGTGCAATACCTGTGATCCCATTAATCTTTTGCACCTTCAAAAGATTAACAGT	4355

3556 AATCTCTGTGCAATCTGTGATCCCATATTCTCTTTGCAACCTCAAAAGATAACAGT 3615
4356 GAGAAATCAGATGCTCTTGTCTCAACTTTTGACCCCTGGCAGGTGATACAGACTGTTCTG 4415
3616 GAGAAATCAGATGCTCTTGTCTCAACTTTTGACCCCTGGCAGGTGATACAGACTGTTCTG 3675
4416 GCCAATATCAGGTATATATATATAGACATGCTAAAGGAGAAACCCCAATCTCTATTTTC 4475
3676 GCCAATATCAGGTATATATATATAGACATGCTAAAGGAGAAACCCCAATCTCTATTTTC 3735
4476 TTTAATCAAGGAATGGGAGATTTCTACTTTGGAGTTCCATCTCAGATTTTGGCGTT 4535
3736 TTTAATCAAGGAATGGGAGATTTCTACTTTGGAGTTCCATCTCAGATTTTGGCGTT 3795
4536 GAGTCTTCGATTAATATTTCAAAATAGCATCAAGGAATTTGTTATTTCTCTTTGCCACAACA 4595
3796 GAGTCTTCGATTAATATTTCAAAATAGCATCAAGGAATTTGTTATTTCTCTTTGCCACAACA 3855
4596 ATTTATAGATTTGGATTTGAAGTGCACCTGATGAAGGATCTCTCGAGTCCCATGCTG 4655
3856 ATTTATAGATTTGGATTTGAAGTGCACCTGATGAAGGATCTCTCGAGTCCCATGCTG 3915
4656 ACCTGGAGCAGCTGGCTTTCTACTATCAGGCAATTTGAAATCTATTTGGAGATGAAGGA 4715
3916 ACCTGGAGCAGCTGGCTTTCTACTATCAGGCAATTTGAAATCTATTTGGAGATGAAGGA 3975
4716 AAACCTCTGTTTGGAGCATTCAAAATAGGAGCATTAATGTTCTGAAAGCATTAATGCG 4775
3976 AAACCTCTGTTTGGAGCATTCAAAATAGGAGCATTAATGTTCTGAAAGCATTAATGCG 4035
4776 TTTGAGTTCAGCAGAGATTTACCTGCTCAGGTCCTGATACAGAAACATCTGGTTCGT 4835
4036 TTTGAGTTCAGCAGAGATTTACCTGCTCAGGTCCTGATACAGAAACATCTGGTTCGT 4095
4836 CTTCATCAGTTGTTCTTCTTAACATAAATCAGAGATACACCATGCTCTGTCATATA 4895
4096 CTTCATCAGTTGTTCTTCTTAACATAAATCAGAGATACACCATGCTCTGTCATATA 4155
4896 GATCTGTTTCATGTTTGGTGGTCTGTTGATGATTTCCCATCTCTGTTGGATGAC 4955
4156 GATCTGTTTCATGTTTGGTGGTCTGTTGATGATTTCCCATCTCTGTTGGATGAC 4215
4956 CCGTTGATCTGACGCTCTTTCAGTATGTTCTTCTATTAACCATCTTATCTCTTCAT 5015
4216 CCGTTGATCTGACGCTCTTTCAGTATGTTCTTCTATTAACCATCTTATCTCTTCAT 4275
5016 TTGATCACCATGGCACATGCTTCAGATACCTACTTACAGTAGACACAGGCTACCCCTT 5075
4276 TTGATCACCATGGCACATGCTTCAGATACCTACTTACAGTAGACACAGGCTACCCCTT 4335
5076 GCTCAGGTTCAAGAGACAGTGAAGGCTCATTTCCGATCTCTTCTTTTTCGAGAAAT 5135
4336 GCTCAGGTTCAAGAGACAGTGAAGGCTCATTTCCGATCTCTTCTTTTTCGAGAAAT 4395
5136 TCTCAATATACAGTGGCTCATTGGGTGATATTTCTGCTGATATTTGGGTCTCA 5195
4396 TCTCAATATACAGTGGCTCATTGGGTGATATTTCTGCTGATATTTGGGTCTCA 4455
5196 CTGAAGATGATCATACCCCTTATCTTCTGCTGCTGATTTGTTTTCACATTTACTT 5255
4456 CTGAAGATGATCATACCCCTTATCTTCTGCTGCTGATTTGTTTTCACATTTACTT 4515
5256 GGGGTAACTCCGCTGAGCACTGCATACCAATTTCTGAGAGGAGATGACAGTGCATC 5315
4516 GGGGTAACTCCGCTGAGCACTGCATACCAATTTCTGAGAGGAGATGACAGTGCATC 4575
5316 TGTAGTATCTATCTTTACCTACAAATTTGTTCTGCTCTTCCAGGAATATTTGGATACT 5375
4576 TGTAGTATCTATCTTTACCTACAAATTTGTTCTGCTCTTCCAGGAATATTTGGATACT 4635
5376 GTAGGCCCTTGTCTCAGAGTGGTGTGATGATCTGCTTACTAACTGTTTGAAGCA 5435
4636 GTAGGCCCTTGTCTCAGAGGCGGTGTGATGATCTGCTTACTAACTGTTTGAAGCA 4695

QY 5436 AAAAACAACCGTGGTCAGGTACCTTAGAAAAAGAAATAGTTTGTATAGAGTTCTCTCATGAC 5495
Db 4696 AAAAACAACCGTGGTCAGGTACCTTAGAAAAAGAAATAGTTTGTATAGAGTTCTCTCATGAC 4755
QY 5496 TATAGTCTGCTCTCGAATCAAGCTTCTCATTTCAAGTGGCCACGCTCTGCAGATCATGAG 5555
Db 4756 TATAGTCTGCTCTCGAATCAAGCTTCTCATTTCAAGTGGCCACGCTCTGCAGATCATGAG 4815
QY 5556 CGAAAGCATCTCTGCTCTCTGCTTTTCTGTTGGGGCTATACCTATATGTTCTCAGAAACATTTGC 5615
Db 4816 CGAAAGCATCTCTGCTCTCTGCTTTTCTGTTGGGGCTATACCTATATGTTCTCAGAAACATTTGC 4875
QY 5616 TGCCAGGAAATTTGTGAACGGGAGAGAGGTTGGAGCTTGCAATTTTTCACGCACTTCACTGT 5675
Db 4876 TGCCAGGAAATTTGTGAACGGGAGAGAGGTTGGAGCTTGCAATTTTTCACGCACTTCACTGT 4935
QY 5676 GGAGCGGAGTCTGCATTTTCTTAAATCAGAGAAATGCGAGTGGTCTCTGTTCAAGGT 5735
Db 4936 GGAGCGGAGTCTGCATTTTCTTAAATCAGAGAAATGCGAGTGGTCTCTGTTCAAGGT 4995
QY 5736 AAAGCCAGAGCTGTGCTCTATCCAGCTCTTACTTTGATGAATATGAGAAACACAGCCCT 5795
Db 4996 AAAGCCAGAGCTGTGCTCTATCCAGCTCTTACTTTGATGAATATGAGAAACACAGCCCT 5055
QY 5796 GGCCTGAAGAGGGGCAACCCCTTCTCATTTATCTCTGAGCGGTATCGAAAGCTCCATTTG 5855
Db 5056 GGCCTGAAGAGGGGCAACCCCTTCTCATTTATCTCTGAGCGGTATCGAAAGCTCCATTTG 5115
QY 5856 GTCTGGCAACCACTGCTATATAGAGAGATTTGCTAGGAGCAAGACTAATCAGATG 5915
Db 5116 GTCTGGCAACCACTGCTATATAGAGAGATTTGCTAGGAGCAAGACTAATCAGATG 5175
QY 5916 TTATTTGATTTCAACTGGCAGTTACTGTGA 5945
Db 5176 TTATTTGATTTCAACTGGCAGTTACTGTGA 5205

RESULT 6

ADS86880

ID ADS86880 standard; cDNA; 5205 BP.

AC ADS86880;

XX 16-DEC-2004 (first entry)

XX cDNA encoding human E3alpha ubiquitin ligase, with SNP.

XX Human; E3alpha ubiquitin ligase; huE3alpha1;

XX ubiquitin-proteasome pathway; chromosome 15 q; rapid muscle wasting;
XX fasting; metabolic acidosis; muscle degeneration; kidney failure;
XX renal cachexia; uremia; diabetes mellitus; sepsis; AIDS wasting syndrome;
XX cancer cachexia; negative nitrogen balance; burn; Cushing's syndrome;
XX inflammatory cachexia; hyperthyroidism; denervation atrophy;
XX protein/tissue wasting; energy-protein malnutrition; muscle atrophy;
XX gene therapy; ss; gene; SNP; single nucleotide polymorphism.

OS Homo sapiens.

XX Key Location/Qualifiers

FH CDS 1..5205

FT /*tag= a

FT /product= "E3alpha ubiquitin ligase"

XX US2004185037-A1.

XX 23-SEP-2004.

XX 15-JAN-2004; 2004US-00758672.

XX 08-MAR-2000; 2000US-0187911P.

XX 28-NOV-2000; 2000US-00724126.

XX

PA (HANH/) HAN H.
XX (KWAK/) KWAK K.
PI Han H, Kwak K;
XX WPI; 2004-707854/69.
XX Novel isolated human E3alpha ubiquitin ligase nucleic acid molecule
PI useful for treating and/or preventing renal cachexia or inflammatory
XX cachexia.
PS Example 9; SEQ ID NO 18; 115pp; English.
XX
XX The present invention relates to new orthologue of human E3alpha
CC ubiquitin ligase, huE3alpha1 and huE3alpha1. Most intracellular proteins
CC are degraded through the ubiquitin-proteasome pathway. Proteins are
CC marked for proteasomal degradation by conjugation of ubiquitin to the
CC protein. Conjugation of the ubiquitin molecule involves the activation by
CC E1 enzyme, subsequent transfer to E2 enzyme, which serves as a carrier,
CC and then interacts with a specific E3 ligase family member. E3 ligase
CC binds to proteins targeted for degradation and catalyses the transfer of
CC ubiquitin from the E2 carrier enzyme to the target protein. E3 ligase
CC determines the specificity of the system. The E3alpha family is the main
CC family of intracellular ligases and is involved in the N-end rule pathway
CC of protein degradation. E3alpha enzyme binds directly to the primary
CC destabilising N-terminal amino acid and catalyses ubiquitin conjugation
CC thereby targeting the protein for degradation. The human E3alpha gene is
CC located on chromosome 15 q. Increased proteolysis through the ubiquitin-
CC proteasome pathway has been determined to be a major cause of rapid
CC muscle wasting including, fasting, metabolic acidosis, muscle
CC degeneration, kidney failure, renal cachexia, uremia, diabetes mellitus,
CC sepsis, AIDS wasting syndrome, cancer cachexia, negative nitrogen
CC balance, burns, Cushing's syndrome, inflammatory cachexia,
CC hyperthyroidism, denervation atrophy, protein/tissue wasting, and energy-
CC protein malnutrition. E3alpha plays a role in the overall increase in
CC ubiquitination that is associated with and may mediate muscle atrophy in
CC catabolic and other disease states. Treatment may be administered by gene
CC therapy, cell therapy and antisense therapy methods. The present sequence
CC is an assembled sequence of human E3alpha ubiquitin ligase. The patent
CC describes this sequence as having a SNP at nucleotide 5397 of the
CC sequence causing a change of cytosine to a thymidine at position 4702 of
CC human E3alpha1. However, this sequence is shown in the sequence listing
XX to have an adenine at position 4702, so no features could be done.
XX
SQ Sequence 5205 BP; 1553 A; 1017 C; 1171 G; 1450 T; 0 U; 14 Other;
Query Match 81.4%; Score 5136; DB 13; Length 5205;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 5189; Conservative 10; Mismatches 6; Indels 45; Gaps 1;
QY 696 ATGCGGACGAGGAGGCTGGAGGTACTGAGAGGATGGAAATCAGCGCGAGGTACCCGAG 755
DB 1 ATGCGGACGAGGAGGCTGGAGGTACTGAGAGGATGGAAATCAGCGCGAGGTACCCGAG 60
QY 756 ACCCTCAGCGCTGGGATCTTGGTGGGATCAGCAAGTTGATTTTATCTGCTTTCTTG 815
DB 61 ACCCTCAGCGCTGGGATCTTGGTGGGATCAGCAAGTTGATTTTATCTGCTTTCTTG 120
QY 816 CATCATTTGGGCAAAATGCTGCGAGAAATTTACTTTGCTGAAATGGACCCAGACTTGGAA 875
DB 121 CATCATTTGGGCAAAATGCTGCGAGAAATTTACTTTGCTGAAATGGACCCAGACTTGGAA 180
QY 876 AAGCAGGAGGAAAGTGTCAAAATGTCATATTTCACTCCACTCGGAATGGTACTTATTTGGA 935
DB 181 AAGCAGGAGGAAAGTGTCAAAATGTCATATTTCACTCCACTCGGAATGGTACTTATTTGGA 240
QY 936 GAAGATCCAGATATTTGCTTAGAGAAATTTAGACACAGTGGGACATTTTCAGCTTTGCGG 995
DB 241 GAAGATCCAGATATTTGCTTAGAGAAATTTAGACACAGTGGGACATTTTCAGCTTTGCGG 300
QY 996 AGGGTTTTCAAAGTGGGAGACAAACCTATTCTTTCAGGAGATTTGTCAATTTGATCCAAACA 1055
DB 301 AGGGTTTTCAAAGTGGGAGACAAACCTATTCTTTCAGGAGATTTGTCAATTTGATCCAAACA 360

QY 1056 TGTGTACTCTGTATGGACTGCTTCCAGGACAGTGTTCATAAAATCATCGTTTCAAGATG 1115
DB 361 TGTGTACTCTGTATGGACTGCTTCCAGGACAGTGTTCATAAAATCATCGTTTCAAGATG 420
QY 1116 CATACTTCTACTGAGGAGGGTCTCTGTGACTGTGTGGAGACACAGAGGCGATGGAAATCGGC 1175
DB 421 CATACTTCTACTGAGGAGGGTCTCTGTGACTGTGTGGAGACACAGAGGCGATGGAAATCGGC 480
QY 1176 CTTTTTTGTGTAAATCATGAACTGTGGAAGAGCGGTACTATATAAAGAGAATTCACGCTGT 1235
DB 481 CTTTTTTGTGTAAATCATGAACTGTGGAAGAGCGGTACTATATAAAGAGAATTCACGCTGT 540
QY 1236 CCGTTGTAATGAGAGGTAAATTTGTCRAAGCCAGGAAATATTTCTTCAGTGATATAATAT 1295
DB 541 CCGTTGTAATGAGAGGTAAATTTGTCRAAGCCAGGAAATATTTCTTCAGTGATATAATAT 600
QY 1296 GTCTGTAGAAATGACTATATGGAAGAGGAAAGAACTGCTCTCTGAACTCCAGATAAGG 1355
DB 601 GTCTGTAGAAATGACTATATGGAAGAGGAAAGAACTGCTCTCTGAACTCCAGATAAGG 660
QY 1356 GAGAAAAATGAAAGATACTATTTGTGTCTTTTCAANTGATGAACACCATTCATATGACAC 1415
DB 661 KNR-----YTCVNDHHSYDH 675
QY 1416 GTCATATACAGCCTACAAAGAGCTCTTGACTGTGAGCTCGCAGAGGCCAGTTCATATACC 1475
DB 676 GTCATATACAGCCTACAAAGAGCTCTTGACTGTGAGCTCGCAGAGGCCAGTTCATATACC 735
QY 1476 ACTGCCATTGCAAAAGAGGGTGTGCGGGCTGTTTAAAGCGGAGCTTATGCTGCTTGCAG 1535
DB 736 ACTGCCATTGCAAAAGAGGGTGTGCGGGCTGTTTAAAGCGGAGCTTATGCTGCTTGCAG 795
QY 1536 GAAGCAAAAGGAGATATAAGAGAGTCATTCAGAAATGTCTCTCAACATCCACTTCATGTA 1595
DB 796 GAAGCAAAAGGAGATATAAGAGAGTCATTCAGAAATGTCTCTCAACATCCACTTCATGTA 855
QY 1596 GAAGTATTACTACTCAGAGATTTAGGCTCATCAGAAATTTGCTTGGCTTCTGTTCTTCTGG 1655
DB 856 GAAGTATTACTACTCAGAGATTTAGGCTCATCAGAAATTTGCTTGGCTTCTGTTCTTCTGG 915
QY 1656 ATGAAACAAAATTAATGAGCTATTCAAGTGACTTTTAGGCGAGATCTTTTGCACAGCATGCC 1715
DB 916 ATGAAACAAAATTAATGAGCTATTCAAGTGACTTTTAGGCGAGATCTTTTGCACAGCATGCC 975
QY 1716 AGAGAGAACCTGACTCGGAGAAATCCCTGTCTCATAGAGAGGTTAATGCTTTGGATGCA 1775
DB 976 AGAGAGAACCTGACTCGGAGAAATCCCTGTCTCATAGAGAGGTTAATGCTTTGGATGCA 1035
QY 1776 AAGCTTTTATAAAGGTGCGGTAAGATCTTTCATCAATTTGATCTTCAGCAGTTTTTTTATG 1835
DB 1036 AAGCTTTTATAAAGGTGCGGTAAGATCTTTCATCAATTTGATCTTCAGCAGTTTTTTTATG 1095
QY 1836 GAGATGGAATACAAAAAACTCTTTGCTATGGAATTTGTGAAAGTATTAATAAACAATCGAG 1895
DB 1096 GAGATGGAATACAAAAAACTCTTTGCTATGGAATTTGTGAAAGTATTAATAAACAATCGAG 1155
QY 1896 AAAGAAATATATCAGTGATGATGACAGAAAGTATCTCTATAAATGCGCATTTTCAGTTTCCAG 1955
DB 1156 AAAGAAATATATCAGTGATGATGACAGAAAGTATCTCTATAAATGCGCATTTTCAGTTTCCAG 1215
QY 1956 ATGTTTTACTGTTCTTACTCTGCGCTCGACATCTTATTTGAAGAGCAGAAATGTTATCTCTGTC 2015
DB 1216 ATGTTTTACTGTTCTTACTCTGCGCTCGACATCTTATTTGAAGAGCAGAAATGTTATCTCTGTC 1275
QY 2016 ATTAAGTGAATACAAAAAACTCTTTGCTATGGAATTTGTGAAAGTATTAATAAACAATCGAG 2075
DB 1276 ATTAAGTGAATACAAAAAACTCTTTGCTATGGAATTTGTGAAAGTATTAATAAACAATCGAG 1335
QY 2076 TTCCAGGGTTATAGCCAGGACAAATTTGGAAAGAGTATATGCAAGTAAATATGACCTTAAG 2135
DB 1336 TTCCAGGGTTATAGCCAGGACAAATTTGGAAAGAGTATATGCAAGTAAATATGACCTTAAG 1395

QY	2136	TATATCTGATCAGCAAAACCCCAATATGACAGAAAGATTAAAGATGCGAGTTCCTTGAA	2195
DB	1396	TATATCTGATCAGCAAAACCCCAATATGACAGAAAGATTAAAGATGCGAGTTCCTTGAA	1455
QY	2196	GGTTTTCGATCTTTTTCGAAGATCTTACCTGTATGACAGGGAATGGAAGAAATCCGAAGA	2255
DB	1456	GGTTTTCGATCTTTTTCGAAGATCTTACCTGTATGACAGGGAATGGAAGAAATCCGAAGA	1515
QY	2256	CAGGTGCGGCAACACATTTGAAGTGGATCCTGTATGCGGAGGCTGCCATTTGCTATACAGATG	2315
DB	1516	CAGGTGCGGCAACACATTTGAAGTGGATCCTGTATGCGGAGGCTGCCATTTGCTATACAGATG	1575
QY	2316	CAATTTGAAGATATTTTACTCATGTTCCAAAGTGGTGTGCTTGTGATGGAAGAACTCTTA	2375
DB	1576	CAATTTGAAGATATTTTACTCATGTTCCAAAGTGGTGTGCTTGTGATGGAAGAACTCTTA	1635
QY	2376	CTTGTGGCTTATAAAGATGTGACAAAGCTGTGATGAGTGCAGTACAGATTTTCATATCT	2435
DB	1636	CTTGTGGCTTATAAAGATGTGACAAAGCTGTGATGAGTGCAGTACAGATTTTCATATCT	1695
QY	2436	AGTAGCAAGACAGTAGTACAATCGTGTGGACATAGTTTGGAAACAAAGTCTCTACAGATTA	2495
DB	1696	AGTAGCAAGACAGTAGTACAATCGTGTGGACATAGTTTGGAAACAAAGTCTCTACAGATTA	1755
QY	2496	TCTGAGGATCTGTGAAGCATACATCTGCCACTCTCTAGACCCCTGCTGCTTCATGTA	2555
DB	1756	TCTGAGGATCTGTGAAGCATACATCTGCCACTCTCTAGACCCCTGCTGCTTCATGTA	1815
QY	2556	CGTTTAAACGAGCTGGTGTGTTTCAAGACTGCAATGAATTTGTCTTTTGGAGACTTT	2615
DB	1816	CGTTTAAACGAGCTGGTGTGTTTCAAGACTGCAATGAATTTGTCTTTTGGAGACTTT	1875
QY	2616	CAAGTAGAGGTACTAGTGGATATCTTTTACGTTCTGTGTGTTGGTTGCCAGGTTGTT	2675
DB	1876	CAAGTAGAGGTACTAGTGGATATCTTTTACGTTCTGTGTGTTGGTTGCCAGGTTGTT	1935
QY	2676	GCTGAGATGTGGCGAAGAAATGGAAGTGTCTTTATAGCCAGGTGTTTATTAACAAGAT	2735
DB	1936	GCTGAGATGTGGCGAAGAAATGGAAGTGTCTTTATAGCCAGGTGTTTATTAACAAGAT	1995
QY	2736	GTTAAGTCAGAGAAATGATGATGAAGATATCATGCTTCAGATTTGGTGGTGCATCT	2795
DB	1996	GTTAAGTCAGAGAAATGATGATGAAGATATCATGCTTCAGATTTGGTGGTGCATCT	2055
QY	2796	TTAATGGATCCCAATAGTCTTGTGTACTGTGTACTTCAGAGGTATGAACCTTGGCGAGCT	2855
DB	2056	TTAATGGATCCCAATAGTCTTGTGTACTGTGTACTTCAGAGGTATGAACCTTGGCGAGCT	2115
QY	2856	TTTAAACAAGACCATATCTACAAAGACAGAGTTTGAATTAACAATATACTAATA	2915
DB	2116	TTTAAACAAGACCATATCTACAAAGACAGAGTTTGAATTAACAATATACTAATA	2175
QY	2916	GAGAAATGCTTCAGGTCTCATCTATATTTGGGTGAGCGTTATGTACCTGGAGTGGGA	2975
DB	2176	GAGAAATGCTTCAGGTCTCATCTATATTTGGGTGAGCGTTATGTACCTGGAGTGGGA	2235
QY	2976	AATGTGACCAAGAGAGGTCAATGAGAGAAATCATTTCACTTGTCTTGCATTTGAACCC	3035
DB	2236	AATGTGACCAAGAGAGGTCAATGAGAGAAATCATTTCACTTGTCTTGCATTTGAACCC	2295
QY	3036	ATGCCACACAGTGCCATTTGCCAAAATTTAATCTGAGAAATGAATAATTAATGAACCTGCTTA	3095
DB	2296	ATGCCACACAGTGCCATTTGCCAAAATTTAATCTGAGAAATGAATAATTAATGAACCTGCTTA	2355
QY	3096	GAGATGTCTAATAAAGTGGCCACATTTAAGAACCCAGGTGTATCAGGCCATGGAGTT	3155
DB	2356	GAGATGTCTAATAAAGTGGCCACATTTAAGAACCCAGGTGTATCAGGCCATGGAGTT	2415
QY	3156	TATGAATTAAGAGATCACTGAAGACTTCAATATGTACTTTTATCATTTACTCCAAA	3215
DB	2416	TATGAATTAAGAGATCACTGAAGACTTCAATATGTACTTTTATCATTTACTCCAAA	2475
QY	3216	ACCCAGCATAGCAGGCTGAACATATGCAAGAAAGAGAGAAACCAAGAAACAAAGAT	3275

DB	2476	ACCCAGCATAGCAGGCTGAACATATGCAAGAAAGAGAGAAACCAAGAAACAAAGAT	2535
QY	3276	GAGCATTTGCCCGCCACACACCTCTCTGAAATTTCTGCCCTCTGCTTTCAGCAAAAGTGAATTAAC	3335
DB	2536	GAGCATTTGCCCGCCACACACCTCTCTGAAATTTCTGCCCTCTGCTTTCAGCAAAAGTGAATTAAC	2595
QY	3336	CTTCTCAATCTGTGATATCATGATGTACATTTCTCAGGACCGTATTTTGAAGCGGCAATAGAC	3395
DB	2596	CTTCTCAATCTGTGATATCATGATGTACATTTCTCAGGACCGTATTTTGAAGCGGCAATAGAC	2655
QY	3396	ACAGATTTCTAATCTGTGACCGAGGAGTCTCCAAATGGCTTTTTCATATTTCTGCGATTTG	3455
DB	2656	ACAGATTTCTAATCTGTGACCGAGGAGTCTCCAAATGGCTTTTTCATATTTCTGCGATTTG	2715
QY	3456	GGTTTACTAGAGAGAGCAACAGCTTCAAAAAGCTCTCTGAAGAAAGATTAACATTTTGAC	3515
DB	2716	GGTTTACTAGAGAGAGCAACAGCTTCAAAAAGCTCTCTGAAGAAAGATTAACATTTTGAC	2775
QY	3516	TTTTATCATAAAGCTTTCAAGATTTGGAGTTTCCAGCATGAATATACAAATGCTTTTGGAA	3575
DB	2776	TTTTATCATAAAGCTTTCAAGATTTGGAGTTTCCAGCATGAATATACAAATGCTTTTGGAA	2835
QY	3576	AAACTCAAGGAAATTTCCCGAGTTAGAGGCGCAGAGGACATGATTAACCTGATCTTCTCAG	3635
DB	2836	AAACTCAAGGAAATTTCCCGAGTTAGAGGCGCAGAGGACATGATTAACCTGATCTTCTCAG	2895
QY	3636	ATGTTTGACACAGTGAAGCGATTAAGAGAAATTTCTGTTTAAATTTGTAGCAACCAATCA	3695
DB	2896	ATGTTTGACACAGTGAAGCGATTAAGAGAAATTTCTGTTTAAATTTGTAGCAACCAATCA	2955
QY	3696	GGATCGGAATCTATTAAGAATGATGAGATTAATCATGATTAAGAAAGAGCAAGCAAGCA	3755
DB	2956	GGATCGGAATCTATTAAGAATGATGAGATTAATCATGATTAAGAAAGAGCAAGCAAGCA	3015
QY	3756	AGAAAGCTGAAGCTGTAGGCTACATCGCCAGAGGATCATGGCTCAGATGCTGCTCTTA	3815
DB	3016	AGAAAGCTGAAGCTGTAGGCTACATCGCCAGAGGATCATGGCTCAGATGCTGCTCTTA	3075
QY	3816	CAGAAAACTTCATTTGAAACTCATAACTCATGTATGACAAATACATCAGAAATGCTCTGG	3875
DB	3076	CAGAAAACTTCATTTGAAACTCATAACTCATGTATGACAAATACATCAGAAATGCTCTGG	3135
QY	3876	AAAGAGATTCATTTAGAGGAGAGAGACCCAGCAGCTCAGTGACTACTCTAGAAAT	3935
DB	3136	AAAGAGATTCATTTAGAGGAGAGAGACCCAGCAGCTCAGTGACTACTCTAGAAAT	3195
QY	3936	GCTTTGGGTCTTAAACCGGGTCCATCTGTTACTGAAAAGGAGGTGCTGACGTGCTCTT	3995
DB	3196	GCTTTGGGTCTTAAACCGGGTCCATCTGTTACTGAAAAGGAGGTGCTGACGTGCTCTT	3255
QY	3996	TGCCAAGAGAAACAGGAGGTGAAAATAGAAAATTAATGCAATGTTATTCGCTCTGCTC	4055
DB	3256	TGCCAAGAGAAACAGGAGGTGAAAATAGAAAATTAATGCAATGTTATTCGCTCTGCTC	3315
QY	4056	CAGAAATCTACTGCTTTAAACCAGCACAGGGGAAAACCCATAGAACTCTCAGGAGAGGCC	4115
DB	3316	CAGAAATCTACTGCTTTAAACCAGCACAGGGGAAAACCCATAGAACTCTCAGGAGAGGCC	3375
QY	4116	CTAGACCACTTTTTCATGAGTCCAGACTTGGCATATGGAACCTTATACAGGAGCTGCTG	4175
DB	3376	CTAGACCACTTTTTCATGAGTCCAGACTTGGCATATGGAACCTTATACAGGAGCTGCTG	3435
QY	4176	CATGTATGACGACGCTGCTGCGCAGAGTATTTTGAAGCTGTACAGCTCAGCTCTCAG	4235
DB	3436	CATGTATGACGACGCTGCTGCGCAGAGTATTTTGAAGCTGTACAGCTCAGCTCTCAG	3495
QY	4236	CAGCGCATTCATGTTGACCTTTTTCGACTTGGAAAGTGGAGATATCTTTGCTCTCTTTC	4295
DB	3496	CAGCGCATTCATGTTGACCTTTTTCGACTTGGAAAGTGGAGATATCTTTGCTCTCTTTC	3555
QY	4296	AAATCTCTGTGCAATATCTGTGATCCCATTTATCTTTTGAACTCTCAAAAGATAAACAAGT	4355

Db 3556 AATCTCTGTGCAATACATGTCATCCCATTTATCTCTTTCACACCTCAAAAGATAAACAAGT 3615
Qy 4356 GAGATGACAGATGCTCTTCTCACTTTTGTACCTTGCACCGGTGATACAGACTGTTCTG 4415
Db 3616 GAGAATGACAGATGCTCTTCTCACTTTTGTACCTTGCACCGGTGATACAGACTGTTCTG 3675
Qy 4416 GCCAGATATCAGGTTTAAATATATAGACATGCTAAAGGAGAAAACCCCAATTCCTATTTTC 4475
Db 3676 GCCAGATATCAGGTTTAAATATATAGACATGCTAAAGGAGAAAACCCCAATTCCTATTTTC 3735
Qy 4476 TTAAATCAAGAAATGGAGATCTCTATCTTGGAGTTCCATTCATCTGAGTTTGGCGTT 4535
Db 3736 TTAAATCAAGAAATGGAGATCTCTATCTTGGAGTTCCATTCATCTGAGTTTGGCGTT 3795
Qy 4536 GAGCTTTCGATTAATATTTCAATAGCATCAAGGAATGGTTATCTCTTTCGCCACAACA 4595
Db 3796 GAGCTTTCGATTAATATTTCAATAGCATCAAGGAATGGTTATCTCTTTCGCCACAACA 3855
Qy 4596 ATTTATAGAAATGGATTTGAAATGGCCACTCATGAAAGGGATCTCTGAGTCCCATGCTG 4655
Db 3856 ATTTATAGAAATGGATTTGAAATGGCCACTCATGAAAGGGATCTCTGAGTCCCATGCTG 3915
Qy 4656 ACCTGGAGCACTGCGCTTTCACTATCCAGGCAATTTGAAATCTATTTGGGAGATGAAGGA 4715
Db 3916 ACCTGGAGCACTGCGCTTTCACTATCCAGGCAATTTGAAATCTATTTGGGAGATGAAGGA 3975
Qy 4716 AAACCTCTGTTGGAGCACTTCAAAATAGGAGCATATGCTCTGAAGCATTAATGCAG 4775
Db 3976 AAACCTCTGTTGGAGCACTTCAAAATAGGAGCATATGCTCTGAAGCATTAATGCAG 4035
Qy 4776 TTTCAGGTTGCACAGAGATTAACCTGCTCTCAGGTCCTGATACAGAAACATCTGCTCGT 4835
Db 4036 TTTCAGGTTGCACAGAGATTAACCTGCTCTCAGGTCCTGATACAGAAACATCTGCTCGT 4095
Qy 4836 CTCTCATCAGTTGTTCTTCTTAACATAAAATCAGAAAGATACACCATGCTTCTGTCTATA 4895
Db 4096 CTCTCATCAGTTGTTCTTCTTAACATAAAATCAGAAAGATACACCATGCTTCTGTCTATA 4155
Qy 4896 GATCTGTTTCATGTTTGGTGGTGCTGTGTAGCATTCCTATCCCATCTTGTATGGATGAC 4955
Db 4156 GATCTGTTTCATGTTTGGTGGTGCTGTGTAGCATTCCTATCCCATCTTGTATGGATGAC 4215
Qy 4956 CCGTGTGATCTGCAGCCTTCTCAGTATGTTCTTCTCTATAACCACTTTATCTCTTCCAT 5015
Db 4216 CCGTGTGATCTGCAGCCTTCTCAGTATGTTCTTCTCTATAACCACTTTATCTCTTCCAT 4275
Qy 5016 TTGATCACCATGGCACACATGCTTCAGATACCTACTTACAGTAGACACAGGCTACCCCTT 5075
Db 4276 TTGATCACCATGGCACACATGCTTCAGATACCTACTTACAGTAGACACAGGCTACCCCTT 4335
Qy 5076 GCTCAGGTTCAAGAGACAGTGAAGAGGCTCATTCGGCATCTTCTTCTTTCGAGAAAT 5135
Db 4336 GCTCAGGTTCAAGAGACAGTGAAGAGGCTCATTCGGCATCTTCTTCTTTCGAGAAAT 4395
Qy 5136 TCTCAATATACAAGTGGCTCCATTCGGTGTGATATCTCTGGCTGTGATTTTGGGTCTCA 5195
Db 4396 TCTCAATATACAAGTGGCTCCATTCGGTGTGATATCTCTGGCTGTGATTTTGGGTCTCA 4455
Qy 5196 CTGAAGAAATGGCATCACCCCTTATCTTCGCTGTGCTGATTTGTTTTCACATTTACTT 5255
Db 4456 CTGAAGAAATGGCATCACCCCTTATCTTCGCTGTGCTGATTTGTTTTCACATTTACTT 4515
Qy 5256 GGGTAACTCCGCTGAGAACTGCATACCAATTCGCAAGAGAGATACAGTGCATC 5315
Db 4516 GGGTAACTCCGCTGAGAACTGCATACCAATTCGCAAGAGAGATACAGTGCATC 4575
Qy 5316 TGTAGCTATCTATCTTTACCTACAATTTGTTCTGCTCTTCCAGGAATATTTGGATACT 5375
Db 4576 TGTAGCTATCTATCTTTACCTACAATTTGTTCTGCTCTTCCAGGAATATTTGGATACT 4635
Qy 5376 GTAAGGCCCTTCTCCAGAGGTTGTGTCAGATCTCTGCTTACTAACTGTTTGAAGCAA 5435
Db 4636 GTAAGGCCCTTCTCCAGAGGCGGTGTGTCAGATCTCTGCTTACTAACTGTTTGAAGCAA 4695

Qy 5436 AAAAAACACCGTGGTCAAGTACCTTAGAAAAAGAAATAGTTTGTATAGAGCTTCTCGATGAC 5495
Db 4696 AAAAAACACCGTGGTCAAGTACCTTAGAAAAAGAAATAGTTTGTATAGAGCTTCTCGATGAC 4755
Qy 5496 TATAGCTGCCTCTCGAATCAAGCTTCTCATTTCAAGGTGCCACCGTCTGCAAGATGAG 5555
Db 4756 TATAGCTGCCTCTCGAATCAAGCTTCTCATTTCAAGGTGCCACCGTCTGCAAGATGAG 4815
Qy 5556 CGAAGCATCTGCTGCTCTGCTTCTGTTGGGCTATACCTATGTTCTCAGAAACATTTGC 5615
Db 4816 CGAAGCATCTGCTGCTCTGCTTCTGTTGGGCTATACCTATGTTCTCAGAAACATTTGC 4875
Qy 5616 TGCCAGGAAATTTGAAACGGGGAAGAGTTGGAGCTTGGCATTTTTCACGCACCTTCACTGT 5675
Db 4876 TGCCAGGAAATTTGAAACGGGGAAGAGTTGGAGCTTGGCATTTTTCACGCACCTTCACTGT 4935
Qy 5676 GGAGCCGAGTCTGCAATTTTCTTAAAAATCAGAAATGCGAGTGGTCTCGTTGAAGGT 5735
Db 4936 GGAGCCGAGTCTGCAATTTTCTTAAAAATCAGAAATGCGAGTGGTCTCGTTGAAGGT 4995
Qy 5736 AAAGCCAGAGCTGTGCTATCCAGCTCTTACTTGGATGATATGGAAGAACAGACCT 5795
Db 4996 AAAGCCAGAGCTGTGCTATCCAGCTCTTACTTGGATGATATGGAAGAACAGACCT 5055
Qy 5796 GGCCTGAAGAGGGGCAACCCCTTCTTATCTCTGAGCGGTATCGGAAGCTCCATTTG 5855
Db 5056 GGCCTGAAGAGGGGCAACCCCTTCTTATCTCTGAGCGGTATCGGAAGCTCCATTTG 5115
Qy 5856 GTCTGGCAACAACACTCATTTATAGAGAGATTCCTAGGAGCCAGAGACTTAATCAGATG 5915
Db 5116 GTCTGGCAACAACAACACTCATTTATAGAGAGATTCCTAGGAGCCAGAGACTTAATCAGATG 5175
Qy 5916 TTATTTGGATTCACCTGGCAGTTACTGTGA 5945
Db 5176 TTATTTGGATTCACCTGGCAGTTACTGTGA 5205

RESULT 7

AAV99308
ID AAV99308 standard; cDNA; 6395 BP.

XX AAV99308;

AC AAV99308;

DT 25-MAR-1999 (first entry)

XX cDNA encoding mouse a ubiquitin-protein ligase, Ubr1.

DE Ubiquitin-protein ligase; Ubr1; mouse; ubiquitinylation; degradation;

XX N-end rule pathway; stress-related muscle wasting; inhibitor; screen; ss.

OS Mus sp.

XX Key Location/Qualifiers

XX CDS 115..5388

XX FT /*tag= a

XX FT /product= "Ubr1"

XX US5861312-A.

XX 19-JAN-1999.

XX 02-DEC-1997; 97US-00982956.

XX 02-DEC-1997; 97US-00982956.

XX (CALY) CALIFORNIA INST OF TECHNOLOGY.

XX Kwon YT, Varshavsky A;

XX WPI; 1999-130395/11.

XX P-PSDB; AAW84351.

PT Mouse and human Ubr1 cDNA - useful for producing recombinant Ubr1
XX polypeptides.

PS Claim 1; Col 15-28; 18pp; English.

XX The present sequence encodes a ubiquitin-protein ligase called Ubr1. The
CC Ubr1 enzymes are involved in protein ubiquitinylation and ultimate
CC degradation through the N-end rule pathway and have been linked to stress
CC -related muscle wasting. Recombinant Ubr1 polypeptides can be used to
CC screen for inhibitors of muscle wasting when this is associated with the
CC N-end rule pathway

XX Sequence 6395 BP; 1802 A; 1376 C; 1519 G; 1698 T; 0 U; 0 Other;

Query Match 70.3%; Score 4435.4; DB 2; Length 6395;

Best Local Similarity 86.3%; Pred. No. 0;

Matches 4997; Conservative 0; Mismatches 721; Indels 69; Gaps 6;

QY	590	TTCCGGGCGCGTGAAGAAGTCTCTCCCTGCTC-TCCGACCGGCCACNGGTTTCCGCT	648
DB	8	TTCCGGGCGCGTGAAGAAGTCTCTCCCTGCTCCTGCGGCGGCCACAGGTTTCCGCT	67
QY	649	TGCTCTCTGCGCGGCGTGGCAACTGCGAGCGTCAAGTTTCCCTCAAGATGGCGGACGAGG	708
DB	68	AGCTGGCGCGCGGCGTGGCAACTGCGGCGCTGCTTCCCTTAAGATGGCGGACGAGG	127
QY	709	AGGCTGGAGGTAAGTCAAGAGATGGAATATCAGCGCGAGTTACCCAGACCCCTCAGCGTC	768
DB	128	AGATGGACGCGCGCGAGAGATGGACGTCAGCGCGGCGCTCCCTGGCGCGGACGCGC	187
QY	769	TGGCATCTGGTGGGATCAGCAAGTTGATTTTATCTACTGCTTCTTGTCATCATTTGGCAC	828
DB	188	CGGCATCTGGTGGGATCAGCAAGTTGATTTTATCTACTGCTTCTTGTCATCATTTGGCAC	247
QY	829	AATTCGTGCGCAAAATTTACTTTGCTGAAATGGACCCAGACTTGGAAAGCAGGAGGAAA	888
DB	248	AATTAAGTCCAGAAATTTATTTGCTGAGATGGACCCAGATTTGGAAAGCAGGAGAGA	307
QY	889	GTGTACAAATGTCATATTTCACTCCACTGGAAATGGTACTTTATTTGGAGAAGATCCAGATA	948
DB	308	GTGTACAGATGTCATATTTCACTCTCTTTGGAGTGGTACTTTATTTGGAGAGATCCGATA	367
QY	949	TTTGTCTAGAGAAATTTGAAGCAGTGGAGCAATTTCACTTTTGGGAGGTTTCAAAA	1008
DB	368	TTTGTCTAGAGAAATTTAAACACAGTGGAGCGTTCCAGTTGTGTGGAGGTTTCAAAA	427
QY	1009	GTGGAGACAACTTATTTCTTGGACGGATTTGCAATTTGATCCAACTGTGTACTCTGTA	1068
DB	428	GTGGAGAAACAACTATTTCTTGGAGGATTTGCAATTTGATCCAACTGTGTACTCTGTA	487
QY	1069	TGGACTGCTTCCAGACAGTGTTCATAAAATCATCGTTTACAGATGCATCTTCTACTG	1128
DB	488	TGGACTGCTTCCAAAGTAGTGTTCATAAAACCATCGTTTACAGATGCATCTTCTACTG	547
QY	1129	GAGGAGGTTCTGTGACTGTGGAGACACAGAGGCAATTTGAGCTTTGGGAGGTTTCAAAA	1188
DB	548	GAGGAGGTTCTGTGACTGTGGAGACACAGAGGCGTGGAAAACTGGCCCTTTTGTGTGG	607
QY	1189	ATCATGAACCTGGAGACGCTACTATAAAGAGAAATTCACGCTGTCGGTGGATGAAG	1248
DB	608	ATCATGAACCTGGAGACGCTACTATAAAGAGAGCTTACATTTGCCCATTTGAATGAAG	667
QY	1249	AGGTAATTTGTCACGCCAGGAAAAATTTTCTTTCAGTGATAAAATATGTCGTAGAAATGA	1308
DB	668	AGGTAATTTGTCACGCCAGGAGAAATTTTCTTTCAGTGATAAAATATGTCGTAGAAATGA	727
QY	1309	CTATATGGAGAGAGAAAAAGAACTGCTCTCTGAACTCCAGATAAGGAGAAAAATGAAA	1368
DB	728	CTATATGGAGAGAGAAAAAGAAATTTGCTCTCTGAACTCCAGATAAGGAGAAAAATGAA	787
QY	1369	GATACATTTGTCCTTTTCAATGATGAACCACTTATATGACCACTCATATACAGCC	1428
DB	788	GATACATTTGTCCTTTTCAACGATGAGCACCATTTGATATGATCATGTGATCTACAGTC	847

QY	1429	TACAAAGAGCTCTTGAAGCTGTGAGCTGCGCAGAGGCCAGTTGATACCACTGCAATTGACA	1488
DB	848	TGCAGAGAGCTCTAGATTTGCGAGCTTGCAGAGGCAAGCTGCAACAGCTGCAATCGACA	907
QY	1489	AAGAGGTCGTGCGGCTGTAAAGCGGAGCTTATGCTGCTTGCAGGAAGCAAAAGGAG	1548
DB	908	AAGAGGTCGTGCGGCTGTCAAGCGAGTGTGTATGCCACTTGCAGGAAGCAAAAGGAG	967
QY	1549	ATATAAAGAGTCATTTCAGAAAAATGCTCTCAACATCCACTTTCATGTAGAAATGATTA	1608
DB	968	ATATAAAGAGTCATTTCAGAAAAATGCTCTCAACATCCACTTTCATGTAGAAATGATTA	1027
QY	1609	CAGAGATTTATGCTCATCAGAAATTTGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCT	1668
DB	1028	CCGTGTTTATGCTCATCAGAAATTTGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCT	1087
QY	1669	TGAGCTATTCAAGTGACTTTTAGGCGAGATCTTTTGCACAGCATGCTTTCAGGAAGAC	1728
DB	1088	TGAGCTATTCAAGTGACTTTTAGGCGAGATCTTTTGCACAGCATGCTTTCAGGAAGAC	1147
QY	1729	ACTCGGAAATCCCTGCTCATTAAGCAGGTTAAATGCTTTGGGATGCAAAAGCTTTATAAG	1788
DB	1148	GCTCTGAAATCCCTGCTCATTAAGCAGACTTAATGCTTTGGGATGCAAAAGCTTTATAAG	1207
QY	1789	GTGCGCGTAAGATCTTTCATGAAATTTGATCTTCAGCAGTCTTTTATGAGAGATGGAATA	1848
DB	1208	GTGCGCGTAAGATCTTTCATGAAATTTGATCTTTAGTAGTCTTTTATGAGAGATGGAATA	1267
QY	1849	AAAAACTCTTGTATGAAATTTGAGAGTATTAATAACATGCGAGAAAGATATATCA	1908
DB	1268	AAAAACTCTTGTATGAAATTTGAGAGTATTAATAACATGCGAGAAAGATATATCA	1327
QY	1909	GTGATGATCATGACAGAGATATCTCTATAACTGCACTTTTCAGTTTCAGATGTTTACTGTT	1968
DB	1328	CGACGACCAAGAGAGATCTCTATAACCGCCCTGCTCGGCGAGATGCTCACCGTCC	1387
QY	1969	CTACTCTGCTGCACTCTTATGAGAGCAGAAATGTTATCTCTGCTCAATTTACTGAACTC	2028
DB	1388	CGACTTGGCGCGGCTCTTATGAGAGCAGAAATGTTATTTCTGCTCAATTTACTGAACTC	1447
QY	2029	TGCTAGAGTTTACCTGAGTACTTGGACAGCAATATAATTCACCTTCAGGGTTATA	2088
DB	1448	TGCTAGAGTTTACCTGAGTACTTGGACAGCAATATAATTCACCTTCAGGGTTATA	1507
QY	2089	GCCAGGACAAATTTGGAGAGTATATGAGTAATATGACCTAAAGTATATCTCTGATCA	2148
DB	1508	GCCAGGACAAATTTGGAGAGTATATGAGTAATATGACCTAAAGTATATCTCTGATTA	1567
QY	2149	GCNAACCCCAATATGGACAGAAATTAAGATTCAGTTCTCTTGAAGGTTTTCGATCTT	2208
DB	1568	GCAAGCTGTCTATGAGACAGACGATTAAGAGCGCAGTTCTTGAAGGTTTTCGCTCTT	1627
QY	2209	TTTTGAGATTTCTTACCTGTATGCGGAAATGGAAGAAATCCGAGACAGGTTGGGCAAC	2268
DB	1628	TTCTGAAGATTTCTTACCTGTATGCGGAAATGGAAGAAATCCGAGACAGGTTGGGCAAC	1687
QY	2269	ACATTTGAAGTGGATCTCTGATTTGGAGGCTGCTTGTATACAGATGCAATTTGAAGATA	2328
DB	1688	ACATTTGAAGTGGATCTCTGATTTGGAGGCTGCTTGTATACAGATGCAATTTGAAGATA	1747
QY	2329	TTTTTACTCATGTTTCCAAAGAGTGTGTGCTGTGATGAGAACTCTTACTTGTGGCTTATA	2388
DB	1748	TTTTTACTCATGTTTCCAAAGAGTGTGTGCTGTGATGAGAACTCTTACTTGTGGCTTATA	1807
QY	2389	AAGATGTCAAAAGCTGTGATGAGGTCAGTCCAGTTTTCATATCTAGTAGCAAGACAG	2448
DB	1808	AAGATGTCAAAAGCTGTGATGAGGTCAGTCCAGTTTTCATATCTAGTAGCAAGACAG	1867
QY	2449	TAGTACAACTCTGTGGACATAGTTTGGAAACAAAGTCTCTACAGAGTATCTCAGGATCTTG	2508
DB	1868	TAGTACAACTCTGTGGACATAGTTTGGAAACAAAGTCTCTACAGAGTATCTCAGGATCTTG	1927


```
Db 4088 CCTGGAGCAGGTGTGGTTTCCATCCAGGCAATCGAATACTGTTGGAGATGAGGNA 4147
Qy 4717 AACCTCTGTTGGAGCACTTCAAAATAGCAGCATTAATGGTCTGAAAGCAATTAATGCAAT 4776
Db 4148 AACCTCTATTTGGAGCACTTCAAAATAGCAGCATTAATGGTCTGAAAGCGCTTAATGCAAT 4207
Qy 4777 TTGCAAGTTCACAGAGGATTAACCTGTCCTCAGTCTGATACAGAAACATCTGTTGCTC 4836
Db 4208 TTGCAAGTTCACAGAGGATTAACCTGTCCTCAGTCTGATACAGAAACATCTGTTGCTC 4267
Qy 4837 TTCTATCAGTGTCTTCTTCAATAAATAAGAGATACACCATGCTTCTGCTATAG 4896
Db 4268 TCCTGTCACTTATCTTCTCACTGCAATCAGAAATACACAGGCTTCTGCTGCTG 4327
Qy 4897 ATCTGTTTCATGTTTGGTGGTGTGTTAGCAATCCATCCCTTGTGTAATGGAGAC 4956
Db 4328 ATCTCTTCCATGTTCTGCTGGCGAGTCTTAGCGTTCCTCAATCTTGTATGGAGTACA 4387
Qy 4957 CTGTTGATCTGAGCTTCTTCACTGTTAGTCTTCTTATTAACCACTTATCTTCCAT 5016
Db 4388 CCGTGGATCTGAGCGCTGCGCACTTAGTCTTCTTATTAACCACTTATCTTCCATC 4447
Qy 5017 TGATCACCATGACACATGCTTCAAGTACTTCTACAGTACACA-----CAG 5064
Db 4448 TGATCACCATGACACATGCTTCAAGTACTTCTACAGTACACA-----CAG 4507
Qy 5065 GCCTACCTCTGCTCAGTTCAGAGAGACAGTGAAGAGGCTCAATCCGATCTCTTCT 5124
Db 4508 GCGCGCGCTTCTGAGGCTGAGAGGATAGTGAAGAGGCTCGCTGTGCACTGCTTCT 4567
Qy 5125 TTGCAGAAATTTCTCAATATACAGTGGTCTCAATGGGTGATATCTTCTGCTGAT 5184
Db 4568 TTGCGAAGTGTGCGAGACACAGACGGCTCTCACTGGGTGGGTGCTCCCGCTGGTACC 4627
Qy 5185 TGTGGGTCTCACTGAAGAGTGCATCACCCCTTATCTTCTGCTGCTGATTTTCTTCC 5244
Db 4628 TGTGGTCTTCTGAGGAGCGCATCACCCCTTATCTTCTGCTGCTGATTTTCTTCC 4687
Qy 5245 ACTATTTACTTGGGTAACTCGCTCAGAGAACTGCATACCAATTTCTGCAAGAGAGT 5304
Db 4688 ACTATTTACTTGGAGTAGCTCGCTGAGAGTGTGTTGGCCAAATCTCTGTAAGAGAGAT 4747
Qy 5305 ACAGTGCACTCTGATGATCTATCTTCTTACCTAAGAAATTTGTTCTGCTCTTCCAGGAAT 5364
Db 4748 TCAAGTGCACTCTGATGATCTATCTTCTTACCTAAGAAATTTGTTCTGCTCTTCCAGGAAT 4807
Qy 5365 ATGGGATACTGTAAAGGCTTGTCTCAGAGGTGTGTGAGATCCTGCTTACTTAACT 5424
Db 4808 ATGGGATACCTAAGGCTTGTCTCAGAGGTGTGTGAGATCCTGCTTACTTCAAGT 4867
Qy 5425 GTTTGAAGCAAAAACACCGTGTGTGAGTACCTTAGAAGAAATAGTTGTATAGAGC 5484
Db 4868 CTTTGAAGCAAGAGTGTGTGTGAGTACCTTAGAAGAAATAGTTGTATAGAGC 4927
Qy 5485 TTCTGTATGATATAGTGTCTTCTGATACAGTCTTCTTCAATTCAGGTGCGCAGGCTG 5544
Db 4928 TTCTGTAGGACTACAGTGTCTTCTAATCAGGCTTCTCACTTTAGGTGTCCAGGCTG 4987
Qy 5545 CAGATGATGCGAAGCATCTGCTCTGCTTCTGCTGCTTCTGCTGCTTCTGCTTCTC 5604
Db 4988 CAGATGATGCGAAGCATCTGCTCTGCTTCTGCTGCTTCTGCTGCTTCTGCTTCTC 5047
Qy 5605 AGAATTTGTCTGCGCAGGAAATTTGGAACGGGAGAGGTTGGAGCTTGCATTTTTCAG 5664
Db 5048 AGAATTTGTCTGCGCAGGAAATTTGGAATGGGAGAGGTTGGAGCTTGCATTTTTCAG 5107
Qy 5665 CACTTACTGTGAGCGGAGTGTGCTTCTTAAATAATCAGAGAAATGCCAGGTGCTCC 5724
Db 5108 CGCTTCAATGTGGTGTGAGTGTGCTTCTTAAATAATCAGAGAAATGCCAGGTGCTCC 5167
Qy 5725 TGGTTGAAGTAAAGCCAGAGCTGTGCTTATCAGCTTCTTCTTGTGATGAATATGAG 5784
Db 5168 TGGTTGAAGTAAAGCCAGAGCTGTGCTTATCAGCTTCTTCTTGTGATGAATATGAG 5227
```

```
Qy 5785 AAAACAGACCCCTGGCCCTGAAGAGGGGCAACCCCTTCAATTTATCTCGTGAAGCGGTATCGGA 5844
Db 5228 AAAACAGACCCAGGGCTAAAGAGAGGAAACCCCACTTCAATTTATCTCGGAGCGGTATCGGA 5287
Qy 5845 AGCTCCATTTGCTCTGGCAACAACACTGCATTTATAGAGAGATGCTAGGAGCCAGAGA 5904
Db 5288 AGCTGCATTTGCTCTGGCAACAACACTGCATTTATAGAGAGATGCTCGAGCCAGAGA 5347
Qy 5905 CTAATCAGATGTTATTTGGATTCAACTGGCAGTTACTGTGAGCTCCAACTCTGCTCTCAAG 5964
Db 5348 CTAATCAGATGTTATTTGGATTCAACTGGCAGTTACTGTGAGCTCTGCTCTCAAG 5407
Qy 5965 ACAATCAAAATGACGACAGTATGAAAGGCTGATTTCAAAATTTATGGAATACTTCTGAGG 6024
Db 5408 ACAATCATGAGTGACATCAATAATAAGACTGATCTAAATAATCTAGAGAACTTCTCTGAGG 5467
Qy 6025 GCTGGGAAGTATTTGGAGGGTCTTTGCTCCATGTCAGAGTTCACTTACATCAATAAAT 6084
Db 5468 ACGGGGGAAGTATTTGGAGGGTCTTTGATCCATGTCAGATTTCAACATTTAATAAAT 5527
Qy 6085 ATTTCTTAATGGAGTATTTGCTTCAATTAGCAAAACATATGCTTCAAGGAAAA-AGGAC 6143
Db 5528 ATTTCTTAATGGAGTATTTGCTTCAATTAGCAAAACATATGCTTCAAGGAAAAACAGAC 5587
Qy 6144 ATAGATCAATCTGT-----TTT 6160
Db 5588 ATAGATTAATGTTTATGTTCTAGAACTAAAGAAATGCTTGTTCATCCAGTGTCTAT 5647
Qy 6161 ATGTGCTAGTATTTCCAGGAATTTATTTCCCTTCAATAATTTGCTCATTTTATTTT 6220
Db 5648 TTCTGCTAATAATTTCCAGAAACTCTTTCCCTTCAATAACTGCTAGTTTCAATTTTCAAT 5707
Qy 6221 CATCACCTTGTGTAGATGAAGTCACTCAAAACAGTTGTAGACATTTATGTTGTGTTAAC 6280
Db 5708 CACCACTGTTTATGAGTCAATTAAGCATTTGTGACATTTCTCCATCTGCTTAAAC 5767
Qy 6281 TCTTCTGCAATTTTGTATTTTGGTGT 6307
Db 5794 ATCTCTGCACTTTTGTATTTGGTGT 5794
```

RESULT 8

AAC86933
ID AAC86933 standard; cDNA; 6395 BP.

AC AAC86933;

XX 02-APR-2001 (first entry)

XX Nucleotide sequence of Mouse Ubr1 protein.

DE Ubr1; E3-type protein; ubiquitin system; ubiquitin-protein ligase;
KW N-end rule pathway; intracellular pathogen; Lysteria monocytogenes;
KW Yersinia enterocolitica; muscle wasting; infection; ss.

XX Mus sp.

XX Key Location/Qualifiers
CDS 115..538
FT /*tag= a
FT /product= "Ubr1"

XX US6159732-A.

XX 12-DEC-2000.

XX 11-JAN-1999; 99US-00228317.

XX 02-DEC-1997; 97US-00982956.

XX (CALY) CALIFORNIA INST OF TECHNOLOGY.

XX

PI Kwon YT, Varshavsky A;
XX WPI; 2001-090278/10.
DR P-PSDB; AAB31162.
XX
PT Inhibiting the N-end rule pathway in mammalian cells for treating
PT infections and various diseases associated with muscle tissue wasting, by
PT inhibiting the expression of Ubri gene.
XX
XX Claim 3; Col 15-28; 18pp; English.
XX
CC The present sequence encodes a murine Ubri enzyme. Ubri is an E3-type
CC protein of the ubiquitin system. Specifically, it is a ubiquitin-protein
CC ligase. The enzyme is specific for destabilizing residues exposed at the
CC N-terminus of protein substrates. Inhibition of the expression of Ubri
CC gene in a cell results in inhibition of the N-end rule pathway. The
CC method is used for treatment of mammalian cells infected with an
CC intracellular pathogen, e.g. *Lysteria monocytogenes* or *Yersinia*
CC *enterocolitica*. Inhibition of N-end rule pathway is also useful for
CC treating various diseases associated with wasting of muscle tissue and
CC infections
XX
SQ Sequence 6395 BP; 1802 A; 1376 C; 1519 G; 1698 T; 0 U; 0 Other;
Query Match 70.3%; Score 4435.4; DB 5; Length 6395;
Best Local Similarity 86.3%; Pred. No. 0;
Matches 4997; Conservative 0; Mismatches 721; Indels 69; Gaps 6;
QY 590 TTCAGGGCCGTCGTAAAGTGTGTCCTCTCTC-TCCGACCGGCCACAGTTTCGCT 648
DB 8 TTCAGGGCCGTCGTAAAGTGTGTCCTCTCTCCTCGCGCCGCCACAGTTTCGCT 67
QY 649 TGCCTCTGCGCGGGGTCCGCACTGACAGGGGTGATTTCCCTCAAGATGCGGACGAG 708
DB 68 AGCTGGCGCGGGGTCCGCACTGCGGGGTTCGTTTCCCTTAAGATGCGGACGAG 127
QY 709 AGGCTGGAGGTACTGAGAGGATGGAATCAGCGCGAGTTACCCAGACCCCTCAGCGTC 768
DB 128 AGATGGACGGCCGAGAGGATGAGCGTCAGCCGAGCCCTCCCTGCGCCCGCAGCGGC 187
QY 769 TGGCATCTTGGTGGGATCAGCAAGTTGATTTTATCTGCTTCTTGTCATCATTTGGCAC 828
DB 188 CGGCATCTGTTGGGATCAGCAAGTTGATTTCTATCTGCTTCTTACATCATTTGGCAC 247
QY 829 AATTGGTCCAGAAATTTACTTGTGTAATGGACCCAGCTTGGAAAGCAGGAGAA 888
DB 248 AATTAGTCCAGAAATTTATTTGCTGAGATGGACCCAGNTTGGAAAGCAGGAGAGA 307
QY 889 GTGTACAAATGTCAATATTTCACTCCACTGGATGGTACTTATTTGGAGAGATCCAGATA 948
DB 308 GTGTACAGATGTCAATATCTACTCTCTTTGGAGTGGTACTTATTTGGAGAGATCCGATA 367
QY 949 TTTGCTTAGAGAAATTTGAAGCAGCTGGAGCAATTTCACTTTTGGGAGGGTTTCAAAA 1008
DB 368 TTTGCTTAGAGAAATTTAAAAACACAGTGGAGCGTTCAGTTGTTGGGAAGGTTTCAAAA 427
QY 1009 GTGGAGAGACAACTATTCTTGGAGGATTTGCAATTTGATTCACATGTGTACTCTGTA 1068
DB 428 GTGGAGAGAACCAATATTCTGTAGGGATTTGCAATTTGATCCACGTTGTGTCTCTGTA 487
QY 1069 TGGACTGCTCCAGGACAGTGTTCATAAAATCATCGTTTACAGATGCATCTTCTACTG 1128
DB 488 TGGACTGCTTCAAAGTAGTGTTCATAAAACCATTCGTTACAGATGCATCTTCTACTG 547
QY 1129 GAGGAGGGTCTGTGACTGTGGAGACACAGAGGCAATGGAAACTGGCCCTTTTGTGTAA 1188
DB 548 GAGGAGGGCTCTGTGACTGTGGAGACACAGAGCGTGGAAACTGGCCCTTTTGTGTGG 607
QY 1189 ATCATGAACCTGGAGAGAGAGTACTATTAAGAGAAATTCAGCTGTTCGTTGAATGAAG 1248
DB 608 ATCAGAGCTGGAAGAGAGAGTACTACAAAGAGAGCTTACATTTGCCCATTTGAATGAAG 667
QY 1249 AGGTAATTTGCCAGCCAGGAAATATTTTCTTCAAGTGAATAATATGTTCGTAGAAATGA 1308

DB 668 AGGTGATTGCTCAAGCCAGGAGAAATATTCCTTTCGGTGAATAAATACATTTGTGAAATGA 727
QY 1309 CTATATGGGAAGAGGAAAAAGAACTGCTCTCTGAACTCCAGATAAGGAGGAAAAAATCAAA 1368
DB 728 CTATATGGGAAGAGGAAAAAGAAATTTGCTCTCTGAACTGCAGATAAGGAGGAAAAAATCAAC 787
QY 1369 GATACATATTGTGCTCTTTTCAATGATGAACACCAATTCATATGACCACTCATATACAGCC 1428
DB 788 GATACATATTGTGCTCTTTTCAACGATGAGCACCATTCGTATGATCATGTGTACAGTC 847
QY 1429 TACAAGAGCTCTTGACTGTGAGCTGCGAGGCCAGTTCGATACCACTGCGCATTTGACA 1488
DB 848 TGCAGAGAGCTCTAGATTGCGAGCTTGCAGAGGCACAGCTGCACACCACTGCCATCGACA 907
QY 1489 AAGAGGGTCTGCGGGCTTTAAAGCGGAGCTTATGCTGCTGCGCAGGAGCAAGAGGAG 1548
DB 908 AAGAGGGTCTGCGGGCTGTCAAAAGCAGGTGTATGCACTTTCGAGAGGAGCAAGAGG 967
QY 1549 ATATAAAGAGTCAATTCAGAAATGTCTCTCAACATCCACTTTCATGTAGAAATTAACA 1608
DB 968 ATATAAAGAGTCACTCAGAGAACGTCTCTCAGCACCCCTCCATGTGAAAGTGTGCACT 1027
QY 1609 CAGAGATTATGGCTCATCAGAAATTTGCTTTCGCTCTTGGTTCCTGATGAACAAAATTA 1668
DB 1028 CCGTGGTTATGGCTCACACAGAAATTCGCTCTCGCGCTTGGCTCTCGATGAACAAAATTA 1087
QY 1669 TGAGCTATTCAAGTGACTTTAGGAGATCTTTTGGCAAGCATGCTTTAGAGAGCAACTG 1728
DB 1088 TGAGCTATTCAAGTGACTTTTAGACAGATAATTTTGGCAGGCTGCTCGTAGAAGAACTG 1147
QY 1729 ACTCGAGAAATCCCTGTCTCATAGAGGTAAATGCTTTTGGGATGCAAGCTTTATATAAG 1788
DB 1148 GCTCTGAATATCCCTGTCTTATAGCAGACTATGCTTTGGGATGCAAACTTTTATAAG 1207
QY 1789 GTGCCCTTAAGATCCTTTCATGAATTTGATCTTCAGCAGTCTTTTATGGAGATGAATACA 1848
DB 1208 GTGCCCTTAAGATCCTTTCATGAATTTGATCTTTAGTAGTCTTTTATGGAGATGAATACA 1267
QY 1849 AAAAACTTTTGTCTATGGAATTTGGAAGATTTATTAACAACTTCAGAGAAATATATCA 1908
DB 1268 AAAAACTTTTGTCTATGGAATTTGGAAGATTTATTAACAACTTCAGAGAGATATATCA 1327
QY 1909 GTGATGATCATGACAGAAATATCTCTATAAATGCTCACTTTTTCAGTTTTCAGATGTTTCTG 1968
DB 1328 GCGAGACACACAGAGAGAGCATCTCCATTAACCGCCCTGTCGTCGAGATGCTCACCGTCC 1387
QY 1969 CTACTCTGGCTCGACATCTTATTTGAAGACAGAAATGTTATCTCTGTCAATTTACTGAACTC 2028
DB 1388 CGACCTTGGCCCGGCATCTTATTTGAAGACAGAGATGTTATTTCTGTCTCAATTTACTGAAACGC 1447
QY 2029 TGCTAGAGATTTTACCTGTGATCTTGGACAGGAAATTAATTCACCTTCCAGGTTTATA 2088
DB 1448 TGCTAGAGATTTTACCTGTGAAATTTGTCAGAGAAATTAATTCACCTTCCAGGTTTATA 1507
QY 2089 GCCAGGACAAATTTGGGAAGAGTATATGACAGTAAATTTGACCTAAAGTATATCTGATCA 2148
DB 1508 GCCAGGACAAATTTGGGAAGAGTCTACGAGTATATGTCAGCTTAAAGTATATCTGATTA 1567
QY 2149 GCAAAACCCCAATATGACAGAAAGATTAAGAAATGCAAGTTCCTTTGAAGTTTTCATCTT 2208
DB 1568 GCAAGCTGTCTATATGACAGAAAGATTAAGAGCGCAGTTCCTGGAAGGTTTCCGGTCTT 1627
QY 2209 TTTTGAAGATTTTACCTGTATGCGGAAATGGAAGAAATCCGAGGACAGGTTGCGCAAC 2268
DB 1628 TTCTGAAGATTTTACCTGTATGCGGAAATGGAAGAAATCAGAGACAGAGTTGGAACAC 1687
QY 2269 ACATTTGAAGTGAATCTGATTTGGGAGGCTGCCATTTGCTATACAGATGCAATTTGAAGATA 2328
DB 1688 ACATTTGAAGTGGACCTGACCTGGAGGCTGCCATGCTATACAGATGCAATTTGAAGATA 1747
QY 2329 TTTTACTCATGTTTCAAGAGGTGTGTGTTGTGATGAAGAACTTACTTGTGCTGCTTATA 2388

Db 1748 TTTTGCTCATGTTCCAGAGTGGTGTGCTTGTGATGAAGATCTCTTACTGTGGCTTATA 1807
Qy 2389 AAGAAATGTCACAAAGCTGTGATGAGGTGCAGTACCAGTTTCATATCTAGTAGCAAGACAG 2448
Db 1808 AAGAAATGTCACAAAGCTGTGATGAGGTGCAGTACCAGTTTCATATCTAGTAGCAAGACAG 1867
Qy 2449 TAGTACAATCGTGTGACATAGTTTGGAAACAAAGTCTCTACAGATATCTGAGATCTTG 2508
Db 1868 TAGTACAATCGTGTGACATAGTTTGGAAACAAAGTCTCTACAGATATCTGAGATCTTG 1927
Qy 2509 TAAGCATATCATCTGCACCTCTCTAGGACCTTGTGCTGTCTCATGTACGTTTTAAAGCAGGC 2568
Db 1928 TAAGCATATCATCTGCACCTCTCTAGGACCTTGTGCTGTCTCATGTACGTTTTAAAGCAGGC 1987
Qy 2569 TGGGTGCTGTTTCAAGATGCAATGAATTTGTCTTTTGGAGCATTTTCAAGTAGAGGTAC 2628
Db 1988 TAGGTGCTGTTTCAAGATGCAATGAATTTGTCTTTTGGAGCATTTTCAAGTAGAGGTAC 2047
Qy 2629 TAGTGGATATCTTTAGTGTCTGTGCTGTGTTGGCTTCCCGAGTGTCTGTGATGTGGC 2688
Db 2048 TGGTGGATATCTTTAGTGTCTGTGCTGTGTTGGCTTCCCGAGTGTCTGTGATGTGGC 2107
Qy 2689 GAAGAAATGGAATGCTCTCTATTAGCCAGGTGTTTATTACCAAGATGTTAAGTGCAGAG 2748
Db 2108 GAAGAAATGGAATGCTCTCTATTAGCCAGGTGTTTATTACCAAGATGTTAAGTGCAGAG 2167
Qy 2749 AAGAAATGATGATAAGATATCATCTGCTTTAGATTTGATGCTTTTAAATGATCCCA 2808
Db 2168 AGGAAATGATGATAAGATATCATCTGCTTTAGATTTGATGCTTTTAAATGATCCCA 2227
Qy 2809 ATAAGTCTTGTGTTACTGTTACTGTTAGAGGTATGAATCTTCCGAGGCTTTTAAAGACCA 2868
Db 2228 ACNAGTCTTGTGTTACTGTTACTGTTAGAGGTATGAATCTTCCGAGGCTTTTAAAGACCA 2287
Qy 2869 TATCTACAAAGACACAGATTTGATTTAAACAATATAATACATAAGTGAAGAAATGCTTC 2928
Db 2288 TATCTACAAAGACACAGATTTGATTTAAACAATATAATACATAAGTGAAGAAATGCTTC 2347
Qy 2929 AGGTCTCTCATCTATATTTGGGTGAGCGTTATGTACCTGGAGTGGAAATGTGACCAAG 2988
Db 2348 AGGTCTCTCATCTATATTTGGGTGAGCGTTATGTACCTGGAGTGGAAATGTGACCAAG 2407
Qy 2989 AAGAGTCAATGAGAGAAATCATCTTGTCTTGTGATTTGAACCCATGCCACACAGTG 3048
Db 2408 AAGAGTCAATGAGAGAAATCATCTTGTCTTGTGATTTGAACCCATGCCACACAGTG 2467
Qy 3049 CCATTTGCCAAATTTTACTGAGATGAATTAATGAATCTGGCTTGTAGAGATGTCTATAA 3108
Db 2468 CCATTTGCCAAATTTTACTGAGATGAATTAATGAATCTGGCTTGTAGAGATGTCTATAA 2527
Qy 3109 ACAAGTGGCCACATTTAAGAAACACAGGTGTATCAGGCCATGGAGTTTATGAACCTTAAAG 3168
Db 2528 ACAAGTGGCCACATTTAAGAAACACAGGTGTGTGGGCCATGGAGTTTATGAATGAAG 2587
Qy 3169 ATGAATCATGAAAGACTTCAATATGTAATTTTATCATTTTATCTTCAAAACCCAGCATAGCA 3228
Db 2588 ATGAATCATGAAAGACTTCAATATGTAATTTTATCATTTTATCTTAAACACACAGCATAGCA 2647
Qy 3229 AGGTGACATATGACAGAAAGAGGAGAAACAGAAACCAAGATCAAGCATTTGGCCG 3288
Db 2648 AGGTGACATATGACAGAAAGAGGAGAAACCAAGATCAAGCATTTGGCCG 2707
Qy 3289 CACCAACCTCTCTGAAATCTGCTGCTCTTTCAGCAAAAGTGAATTAACCTTCTCAACTGTG 3348
Db 2708 CGCCACCTCTCTGAAATCTGCTGCTCTTTCAGCAAAAGTGAATTAACCTTCTCAACTGTG 2767
Qy 3349 ATATCATGATGTACATTTCTAGGACCGTATTGTAGCGGGCAATAGACACAGATTTCTAACT 3408
Db 2768 ATGTTATGATATACATCTCTAGGACCACTTTTGTAGCGGGCAATAGACACAGATTTCTAACT 2827
Qy 3409 TGTGACCGAAGGATGTCTCCAAATGGCTTTTATATTTCTGCGCATTTGGGTTTCTAGAG 3468
Db 2828 TGTGACCGAAGGATGTCTCCAAATGGCTTTTATATTTCTGCGCATTTGGGTTTCTAGAG 2887

Qy 3469 AGAAGCAACAGCTTCAAAAAGCTCTGAAAGAAAGTAACTTTGACTTTTATATCAATAGG 3528
Db 2888 AGAAGCAACAGCTTCAAAAAGCTCTGAAAGAAAGTAACTTTGACTTTTATATCAATAGG 2947
Qy 3529 CTTCAAGATTCGGGAAGTTTCAGCCAT-----GAATATACAAATGCTTTTGGAAAGAC 3579
Db 2948 CTTCAAGATTCGGGAAGTTTCAGCCATGAATGCTCAGAAATATACAAATGCTTTTGGAAAGAC 3007
Qy 3580 TCAAGGAATTTCCCCAGTTTGAAGGCCAGAAAGGACATGATACGTTGGATCTTTTCAGATGT 3639
Db 3008 TCAAGGAATTTCCCCAGTTTGAAGGCCAGAAAGGACATGATACGTTGGATCTTTTCAGATGT 3067
Qy 3640 TTGACACAGTGAAGGATTAAGAGAAATCTTGTTTAATTTGTAGCAACCAATCAGAT 3699
Db 3068 TTGACACAGTGAAGGATTAAGAGAAATCTTGTTTAGTTGTGGCAACCACTTCAGGAC 3127
Qy 3700 CGGAATCTATTAAAGATGATGAGATTTACTCATGATTAAGAAAGAAAGCAAGAAAGAA 3759
Db 3128 TGGAGTGCAATTAAGAGTGAAGGATTTACTCATGATTAAGAAAGAAAGCAAGAAAGAA 3187
Qy 3760 AAGCTGAAGCTGTAGGCTACATCGCCAGAAAGATCATGGCTCAGATGTCTGTCTTACAGA 3819
Db 3188 AAGCTGAAGCTGTAGGCTTCTATCGCCAGAAAGATCATGGCCAGATGTCTGTCTTACAGA 3247
Qy 3820 AAAAATCTTCAATTTGAAAATCTCAATGATGACAATATCATCAAGAAATGCTTGGGAAG 3879
Db 3248 AAAAATCTTCAATTTGAAAATCTCAATGATGACAATATCATCAAGAAATGCTTGGGAAG 3307
Qy 3880 AAGATTCATTTTGGGAGAGAGGACCCAGCAGTCAAGTCACTTCTAGAAATTTGCTT 3939
Db 3308 AAGATTCATTTTGGGAGAGAGGACCCAGCAGTCAAGTCACTTCTAGAAATTTGCTT 3367
Qy 3940 TGGGTCTTAAACGGGCTCCATCTGTACTGAAAAGAGGTGCTGACGTGCATCTCTTGGC 3999
Db 3368 TGGGCTCTTAAACGGGCTCCGCTGTACCGAAAGAGGTGCTGACGTGCATCTCTTGGC 3427
Qy 4000 AAGAAAGACAGAGGTGAAAATAGAAAATATGCAATGATGATTTATCGGCTGTGTCCAGA 4059
Db 3428 AAGAAAGACAGAGGTGAAAATAGAAAATATGCAATGATGATTTATCGGCTGTGTCCAGA 3487
Qy 4060 AATCTACTGCTTAAACCCAGCAGAGGGAAGAAACCATAGAACTCTCAGAGAGAGCCCTAG 4119
Db 3488 AATCCAGCGCTTAAACCCAGCAGAGGGAAGCCCTGTGTGACCACTTAAAGGGAACACCTGG 3547
Qy 4120 ACCCACTTTTCAATGATCCAGACTTGGCATATGGAATTTATACAGAAAGCTGTGTCTATG 4179
Db 3548 ACCCTTTTCAATGATCCAGACTTGGCATATGGAATTTATACAGAAAGCTGTGTCTATG 3607
Qy 4180 TAATGCAAGCTGTGTGGCAGAAAGTATTTTGAAGCTGTACAGCTGAGCTCTCAGAGCAGC 4239
Db 3608 TAATGCAAGCTGTGTGGCAGAAAGTATTTTGAAGCTGTGTGAGCTGTGAGCTGTGAGCAGC 3667
Qy 4240 GCATTCATGTTGACCTTTTGAAGTGTGAAAGTGAAGATATCTTGTGCTCTTGTGCAAT 4299
Db 3668 GCATTCATGTTGACCTTTTGAAGTGTGAAAGTGAAGATATCTTGTGCTCTTGTGCAAT 3727
Qy 4300 CTCTGTGCAATGATGATCCCAATTTCTTTTGAACCTTCAAAAGATAAAGCTGTGAGA 4359
Db 3728 CTCTGTGCAACATGATGATCCCAATTTCTTTTGAACCTTCAAAAGATAAAGCTGTGAGA 3787
Qy 4360 ATGCAAGTGTCTTGTCTCAATTTTGAACCTTGGCAAGTGGATACAGACTGTCTTGTGCA 4419
Db 3788 ATGCAAGTGTCTTGTCTCAATTTTGAACCTTGGCCCGGTGATACAGACTGTCTTGTGCA 3847
Qy 4420 GAATATCAGGTTTAAATATAGACATGCTAAGGAGAAAGCC---CAATTCCTATTTTCT 4476
Db 3848 GAATATCAGGTTTAAATATAGACATGCTAAGGAGAAAGCCCGAGCTGTCTTGTCTTGT 3907
Qy 4477 TTAATCAGGAAATGCGGATTTCTACTTTGGAGTTTCCATTTCCATCTGAGTTTTCGGCTTG 4536
Db 3908 TTAATCAGGAAATGCGGATTTCAACTTTTGTAGTTTCACTTCCATCTGAGTTTGTGGATTC 3967

QY 4537 AGTCTTCGATTAATAATTAATCAATAGCATCAAGGAAATGGTTATTTCTCTTTGCCACAACAA 4596
DB 3968 AGTCTTCGATTAATAATTAATCAATAGCATCAAGGAAATGGTTATTTCTCTTCGCCACAACAA 4027
QY 4597 TTTATAGAAATTTGGAATGGAAGTGCACCTGATGAAAGGATCTCTCGAGTCCCGCATGCTGA 4656
DB 4028 TTTACAGAAATTTGGCTGGAAGTGCCTCTGATGAATAGACCCACGAGTCCCATGATGA 4087
QY 4657 CTTGGAGCAGCTGGCTTTCTACTATCCAGGCAATTTGAAATCTATTGGAGATGAAGAA 4716
DB 4088 CTTGGAGCAGCTGGCTTTCTACTATCCAGGCAATTTGAAATCTATTGGAGATGAAGAA 4147
QY 4717 AACCTCTGTTGGAGCACTTCAAAATAGGAGCAATAATGGTCTGAAAGCAATAATGCAGT 4776
DB 4148 AACCTCTATTGGAGCACTTCAAAATAGGAGCAATAATGGTCTGAAAGCAATAATGCAGT 4207
QY 4777 TTGCAGTTGCAACAGAGGATTAACCTGCTCTAGGTCCTGATACAGAAACATCTGTTGCTC 4836
DB 4208 TTGCAGTTGCAACAGAGGATTAACCTGCTCTAGGTCCTGATACAGAAACATCTGTTGCTC 4267
QY 4837 TTCTATCAGTTGTTCTTCTTAACATATAAATCAGAGATACACCATGCTCTGCTCTATAG 4896
DB 4268 TCCTGTCTAGTTATTTCTTCAACCTGCAATCAGAAATATACACAGGCTCTGCTCTG 4327
QY 4897 ATCTGTTTCATGTTTGGTGGTCTGTTAGCAATCCCATCTCTGTTATTTGGGATGACC 4956
DB 4328 ATCTTCTCCATGTTCTGCTGGCGAGTCTTAGGTTCCCATCTCTGTTATTTGGGATGACA 4387
QY 4957 CTGTTGATCTGAGCCTTTCTTCACTAGTTCTTCTCTATTAACCATCTTATCTTCCAT 5016
DB 4388 CCGTGGATCTGAGCCTGCGCACCTTAGTTCTTCTATATAACCATCTTATCTTCCATC 4447
QY 5017 TGATCACCATGGCACATGCTTCTGATACACTTACAGTAGACA-----CAG 5064
DB 4448 TGATCACCATGGCGCACATGCTTACAGTACTCTTTACACAGATACAGATCTGCTCCAG 4507
QY 5065 GCCTACCCCTTGTCTCAGGTTCAAGAGACAGTGAAGGCTCATTTCCGCACTCTTCTTCT 5124
DB 4508 GCGCGCCCTTGTCTGAGGTTGAGAGGATAGTAGAGGCTCGCTGTCATCTGCTTCT 4567
QY 5125 TTGCAGAAATTTCTCAATATACAGTGGCTCCATTTGGGTGATATTTCTCGCTGGTANT 5184
DB 4568 TTGTGGAAGTGTGCGACACACAGACGCGCTCACTGCGGTGCGGTCTCCCGGCTGGTACC 4627
QY 5185 TGTGGTCTCACTGAAGATGGCATCCCTTATCTTCTGCTGCTGCTGATGTTTTC 5244
DB 4628 TGTGGCTCTCCCTGAGGAAAGGATCACTCCCTTACCTCCGCTGCTGCACTGCTTTTC 4687
QY 5245 ACTATTTACTTTGGGTAACTCCGCTGAGGAACCTGCATACCAATTTCTGCAGAGGAGAGT 5304
DB 4688 ACTATTTACTTTGGAGTAGCTCGCTCTGAAGAACTGTTTGGCCAAATTTCTGCTGAAGAGAA 4747
QY 5305 ACAGTGCACCTGTAGCTATCTATCTTTACCTACAAATTTGTTCTGCTTCTCCAGGAAT 5364
DB 4748 TCAGTGCACCTGTAGCTATCTATCTTTACCCACAAATTTGTTCTGCTTCTCCAGGAAT 4807
QY 5365 ATTGGGATCTGTAAAGCCCTTGTCTCCAGAGGTGTGTGAGATCCTGCTTACTAACT 5424
DB 4808 ATTGGGATACCATAGGCGCTTACTACAGAGGTGTGTGAGATCCTGCTTACTCAAGT 4867
QY 5425 GTTTGAAGCAAAACACCGTGTGCTAGTACCTCCCTAGAAAGAAATAGTTTGAATAGAGC 5484
DB 4868 CTTTGAAGCAGAAAGTGTGTGCTGAGTACCTCCCTAGAAAGAAATAGTTTGAATAGAGC 4927
QY 5485 TTCTGTATGATATAGTCTCTCTGATCAAGCTTCTCATTTACAGGTGCCACCGTCTG 5544
DB 4928 TTCTGTAGGACTACAGCTGTCTCTTAATCAGGCTTCTCACTTTTAGGTGTCCACGGTCTG 4987
QY 5545 CAGATGATGAGGAAAGCATCTGCTCTGCTTTCTGCTGGGCTATACATAGTTCTC 5604
DB 4988 CAGATGATGAGGAAAGCATCTGCTCTCTGCTTTCTGCTGGGCTATCTGTTCTC 5047
QY 5605 AGAACATTTGCTGCCAGGAAATTTGGAACGGGAGAGGTTGGAGCTTGGCATTTTTCACG 5664

DB 5048 AGAACATCTGCTGCTCCAGAAATAGTGAAATGGGGAAGAGGTTGGAGCGTGGCTTTTCATG 5107
QY 5665 CACTTCACTGTGAGCGGAGTCTGCATTTTCTTAAATAACAGAGATGCCAGTGTCC 5724
DB 5108 CGCTTCACTGTGCTGGAGTCTGCATTTTCTTAAATAACCGAGAAATGCAAGGTGTCC 5167
QY 5725 TGGTTGAAGGTAAAGCCAGAGGCTGTGCTTATCCAGCTCTTACTTCTGATGAATATGGAG 5784
DB 5168 TGGTGAAGAAAGCCAGAGGCTGTGCTTACCAGCCCTTACTTGTGATGAATATGGAG 5227
QY 5785 AAACAGACCCCTGCTGGAAGGGGCAACCCCTTCTATTTATCTCTGAGCGGTATCGGA 5844
DB 5228 AAACAGACCCAGGCTTAAAGAGAGAAACCCACTTCACTTATCTCGGAGCGGTATCGGA 5287
QY 5845 AGCTCCATTTGCTGCAACACACTGCATATAGAGAGATGCTTAGGAGCCCAAGGA 5904
DB 5288 AGCTGCATTTGCTGCAACAGCCTTATAGAGAGATGCTCGGAGCCAGGAGA 5347
QY 5905 CTATCAGATGTTATTTGGATTTCACTGGCAGTCTTCTGAGCTCCAACTCTGCTCAAG 5964
DB 5348 CTATCAGATGCTATTTGGATTTAACTGGCAGTCTTCTGAGCTTCACTTCTGCTCAAG 5407
QY 5965 ACAATCAAAATGACGACAGTAGTAAAGGCTGATTTCAAAATTTATGGAATACTTTCTGAGG 6024
DB 5408 ACAATCATGATGACATCAATAATAAAGACTGATCTAAATTTCTAGAGAACTTTCTGAGG 5467
QY 6025 GCTGGGAAAGTATTTGGAGGCTTTTCTGCTCAATGCTCAGGTTCACTTACATCAATAAAT 6084
DB 5468 ACGGGGAAGTATTTGGAGGCTTTTGTGATCCATGTCAGATTTCAACACATTAATAAAT 5527
QY 6085 ATTTCTTAATGGAGTATTTGCTTTCAAATTTAGCAACATATGCTTCAAGGAAAAA-AGGAC 6143
DB 5528 ATTCTTAAATGGAATATTTGCTTTCAATTTCAACATAGCTTCAAGGAAAAAACAAGAC 5587
QY 6144 ATAGATCAATCTGT-----TTT 6160
DB 5588 ATAGATTAATGTTTATGTTCTAGAACACTAAAGAAATGCTTGTTCATCCAAAGTGTCTAT 5647
QY 6161 ATGTGCTAGTATTTCCAGGAATTTATTTCCCTTCAATAATTTGCTCTCATTTTCAATTTATTT 6220
DB 5648 TTCTGCTAATATTTCCAGAAAACTCTTTCCCTTCAATACTGCTTCAATTTTCAATTTTCAAT 5707
QY 6221 CATCCACTTGTAGATGAAGTCACTCAAAACAGTTGTAGACATTTTATGTTGTTGTTAAAC 6280
DB 5708 CACCACCTGTTAAAGGTCACATTAAGCATTTGTGGAATTTTCTCCATCTGGCTAAC 5767
QY 6281 TCTTCTGCAATTTTGTATTTGGGTTTT 6307
DB 5768 ATCTCTGACCTTTGTATTTGGGTTTT 5794

RESULT 9
ADA53348
ID ADA53348 standard; cDNA; 3059 BP.
XX
AC ADA53348;
XX
DT 20-NOV-2003 (first entry)
XX
Human coding sequence, SEQ ID 916.
XX
Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
KW Gene therapy; human; secretory protein; membrane proteins; cancer;
KW inflammatory disease; osteoporosis; neurological disease; gene; ss.
XX
OS Homo sapiens.
XX
PN EP1293569-A2.
XX
PD 19-MAR-2003.
XX
PF 21-MAR-2002; 2002EP-00006586.

[illegible]

Db 1681 TGATGAAGAACTCTTACTTGTGGCTTATPAAAGAAATGTCAAAAGCTGTGATGAGGTGCAG 1740
 Qy 2420 TACAGATTTTATATCTAGTAGCAAGACAGTAGTAGTACAATCGTGTGACATATAGTTTGGAAAC 2479
 Db 1741 TACCAGTTTTCATATCTAGTAGCAAGACAGTAGTAGTACAATCGTGTGACATATAGTTTGGAAAC 1800
 Qy 2480 AAAGTCCCTACAGAGTATCTGAGGATCTTGTAAAGCATATACATCGCCACCTCTCTAGAGCCCT 2539
 Db 1801 AAAGTCCCTACAGAGTATCTGAGGATCTTGTAAAGCATATACATCGCCACCTCTCTAGAGCCCT 1860
 Qy 2540 TGCTGGTCTTTCATGTACGTTTAAAGCAGGCTGGGTGCTGTTTCAAGACATGCAATGTTGT 2599
 Db 1861 TGCTGGTCTTTCATGTACGTTTAAAGCAGGCTGGGTGCTGTTTCAAGACATGCAATGTTGT 1920
 Qy 2600 GTCTTTTGGAGCATCTTCAAGTAGAGGTACTAGTGAATATCCTTTACGTTGTCTGGTGT 2659
 Db 1921 GTCTTTTGGAGCATCTTCAAGTAGAGGTACTAGTGAATATCCTTTACGTTGTCTGGTGT 1980
 Qy 2660 GGTTCGCCAGGTGTGCTGAGATGTGGCGAAGAAATGGACTGTCTCTTATTTAGCCAGGT 2719
 Db 1981 GGTTCGCCAGGTGTGCTGAGATGTGGCGAAGAAATGGACTGTCTCTTATTTAGCCAGGT 2040
 Qy 2720 GTTTTATTTACCAAGATGTTAAAGTCAGAGAGAAATGTATGATAAGATATCATCATGCT 2779
 Db 2041 GTTTTATTTACCAAGATGTTAAAGTCAGAGAGAAATGTATGATAAGATATCATCATGCT 2100
 Qy 2780 TCAGATTGGTGCACTTTTAATGGAATCCCAATGAATGTTTGTACTGTGTAATCTGAGAGTA 2839
 Db 2101 TCAGATTGGTGCACTTTTAATGGAATCCCAATGAATGTTTGTACTGTGTAATCTGAGAGTA 2160
 Qy 2840 TGAATCTGCCAGGCTTTTAAACAGAGCCATATCTACAAAGACAGGATTTGATTAAACA 2899
 Db 2161 TGAATCTGCCAGGCTTTTAAACAGAGCCATATCTACAAAGACAGGATTTGATTAAACA 2220
 Qy 2900 ATATAATACACTAATAGAGAAATGCTTCAGGTCTCTCATCTATATTTGGGTGAGCGTTA 2959
 Db 2221 ATATAATACACTAATAGAGAAATGCTTCAGGTCTCTCATCTATATTTGGGTGAGCGTTA 2280
 Qy 2960 TGTACTGTGAGTGGAAATGTGACAAAGAGAGGTCAATAGAGAGAAATCAATTCACCT 3019
 Db 2281 TGTACTGTGAGTGGAAATGTGACAAAGAGAGGTCAATAGAGAGAAATCAATTCACCT 2340
 Qy 3020 GCTTTGCATTCGACCCATGCCACACAGTGCCTATTCGCAAAATTTACCTGAGAGTGAAGA 3079
 Db 2341 GCTTTGCATTCGACCCATGCCACACAGTGCCTATTCGCAAAATTTACCTGAGAGTGAAGA 2400
 Qy 3080 TAATGAAACTGGCTTAGAGAAATGTATAAACAAAGTGGCCACCATTTAAGAAACAGAGTGT 3139
 Db 2401 TAATGAAACTGGCTTAGAGAAATGTATAAACAAAGTGGCCACCATTTAAGAAACAGAGTGT 2460
 Qy 3140 ATCAGGCCATGGAGTTTATGAATCAAAAGATGAATCACTGAAAGACTTCAATATGTACTT 3199
 Db 2461 ATCAGGCCATGGAGTTTATGAATCAAAAGATGAATCACTGAAAGACTTCAATATGTACTT 2520
 Qy 3200 TTATCATTTACTTCCAAACCCAGCATAGCAAGGCTGAACATATGCAAGAGAAAGAGGAAA 3259
 Db 2521 TTATCATTTACTTCCAAACCCAGCATAGCAAGGCTGAACATATGCAAGAGAAAGAGGAAA 2580
 Qy 3260 ACAAGAAAACAAAGATGAAGCATTTGCCGCCACACCATCTCTGAATTTCTGCCCTGCTTT 3319
 Db 2581 ACAAGAAAACAAAGATGAAGCATTTGCCGCCACACCATCTCTGAATTTCTGCCCTGCTTT 2640
 Qy 3320 CAGCAAGTGAATTAACCTTCACTGATGATATCATGATGTACATTCACGAGACCGTATT 3379
 Db 2641 CAGCAAGTGAATTAACCTTCACTGATGATATCATGATGTACATTCACGAGACCGTATT 2700
 Qy 3380 TGAGCGGCAATAGACACAGATTTCTAACTTGTGACCGAAGGATGCTCCAAATGGCTTT 3439
 Db 2701 TGAGCGGCAATAGACACAGATTTCTAACTTGTGACCGAAGGATGCTCCAAATGGCTTT 2760
 Qy 3440 TCATATTTCTGGCATTTGGGTTTACTAGAGAGAGAGCAACAGCTTCAAAAGCTCTGGAAGA 3499
 Db 2761 TCATATTTCTGGCATTTGGGTTTACTAGAGAGAGAGCAACAGCTTCAAAAGCTCTGGAAGA 2820

Qy 3500 AGAAGTAACATTTGACCTTTTATCATTAAGGCTTCAAGATTCGGAAGTTTTCAGCCATGAATAT 3559
 Db 2821 AGAAGTAACATTTGACCTTTTATCATTAAGGCTTCAAGATTCGGAAGTTTTCAGCCATGAATAT 2880
 Qy 3560 ACAAAATCTTTTGGAAAAAATCTCAAGGAATTTCCCACTTGAAGGCCAGAGGACATGAT 3619
 Db 2881 ACAAAATCTTTTGGAAAAAATCTCAAGGAATTTCCCACTTGAAGGCCAGAGGACATGAT 2940
 Qy 3620 AACGTGATATCTTCAAGATGTTTGAACAGTGAAGCGATTAAGAGAAAAATCTTGTTTAAT 3679
 Db 2941 AACGTGATATCTTCAAGATGTTTGAACAGTGAAGCGATTAAGAGAAAAATCTTGTTTAAT 3000
 Qy 3680 TGTAGCAACACATCAGATCGGAATCTATTAAAGATGATGAGATTACTCATGATAAG 3738
 Db 3001 TGTAGCAACACATCAGATCGGAATCTATTAAAGATGATGAGATTACTCATGATAAG 3059

RESULT 10
 AAH14878
 ID AAH14878 standard; cDNA; 2550 BP.
 XX
 AC AAH14878;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human cDNA sequence SEQ ID NO:12731.
 XX
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 FN EP1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-00116126.
 XX
 PR 29-JUL-1999; 99JP-00248036.
 PR 27-AUG-1999; 99JP-00300253.
 PR 11-JAN-2000; 2000JP-00118776.
 PR 02-MAY-2000; 2000JP-00183767.
 PR 09-JUN-2000; 2000JP-00241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 DR WPI; 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
 PT
 XX
 PS Claim 8; SEQ ID NO 12731; 2537pp + Sequence Listing; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the

CC	detection and/or diagnosis of the abnormality of the proteins encoded by
CC	the full-length cDNAs. The primers allow obtaining of the full-length
CC	cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC	AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC	represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC	oligonucleotides, all of which are used in the exemplification of the
CC	present invention
xx	
SQ	Sequence 2550 BP; 830 A; 454 C; 558 G; 708 T; 0 U; 0 Other;
	Query Match 40.4%; Score 2546.8; DB 4; Length 2550;
	Best Local Similarity 99.9%; Pred. No. 0;
	Matches 2548; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy	1189 ATCATGAACCTGGAGAGCAGGTACTATAAAGAGAAATTCACGCTGCCGTGAATGAAG 1248
Db	1 ATCATGAACCTGGAGAGCAGGTACTATAAAGAGAAATTCACGCTGCCGTGAATGAAG 60
Qy	1249 AGGTAATTGTCCAAGCCAGGAAATATTTCTTCAGTGATATAAATATGTCGTGAGAAATGA 1308
Db	61 AGGTAATTGTCCAAGCCAGGAAATATTTCTTCAGTGATATAAATATGTCGCTGAGAAATGA 120
Qy	1309 CTATATCGGAAGAGGAAAAAGAACTGCTCTCTGAACCTCCAGATAAGGAGGAGAAAAATGAAA 1368
Db	121 CTATATCGGAAGAGGAAAAAGAACTGCTCTCTGAACCTCCAGATAAGGAGGAGAAAAATGAAA 180
Qy	1369 GATACTATTGTGTCCTTTTCAATGATGAACACCAATTCATATGACCACGTCATATACAGCC 1428
Db	181 GATACTATTGTGTCCTTTTCAATGATGAACACCAATTCATATGACCACGTCATATACAGCC 240
Qy	1429 TACAAAGAGCTCTTGACTGTGAGCTCGAGAGGCCAGTTGCAATACACATGCCATTGACA 1488
Db	241 TACAAAGAGCTCTTGACTGTGAGCTCGAGAGGCCAGTTGCAATACACATGCCATTGACA 300
Qy	1489 AAGAGGGTCGTCGGGCTGTTAAAGCGGGAGCTTATGCTGCTGCCAGGAGCAAGGAAG 1548
Db	301 AAGAGGGTCGTCGGGCTGTTAAAGCGGGAGCTTATGCTGCTGCCAGGAGCAAGGAAG 360
Qy	1549 ATATAAGAGCTCATTCAGAAAAATGTCTCTCAACATCCACATCCATGTAGAGATATTACACT 1608
Db	361 ATATAAGAGCTCATTCAGAAAAATGTCTCTCAACATCCACATCCATGTAGAGATATTACACT 420
Qy	1609 CAGAGATTATGGCTCATCAGAAATTTGCTTTGGTCTTGTTCTCTGATGAAACAAATTA 1668
Db	421 CAGAGATTATGGCTCATCAGAAATTTGCTTTGGTCTTGTTCTCTGATGAAACAAATTA 480
Qy	1669 TGAGCTATTCAAGTGACTTTAGCCAGATCTTTTGGCCAGCATGCCCTTAGAGAGACCTTG 1728
Db	481 TGAGCTATTCAAGTGACTTTAGCCAGATCTTTTGGCCAGCATGCCCTTAGAGAGACCTTG 540
Qy	1729 ACTCGAGAGATCCTGTCTCATAGCAGGTTAATGCTTTGGGATGCAAGACTTTATAAAG 1788
Db	541 ACTCGAGAGATCCTGTCTCATAGCAGGTTAATGCTTTGGGATGCAAGACTTTATAAAG 600
Qy	1789 GTGCCCGTAAGATCCTTCATGAATTGATCTTCAGCAGTGTTTTTTATGGAGATGGAATACA 1848
Db	601 GTGCCCGTAAGATCCTTCATGAATTGATCTTCAGCAGTGTTTTTTATGGAGATGGAATACA 660
Qy	1849 AAAAATCTTTGCTATGGAATTTGTGAGTATTATATAAGAACTGCAGAAAGAATATATCA 1908
Db	661 AAAAATCTTTGCTATGGAATTTGTGAGTATTATATAAGAACTGCAGAAAGAATATATCA 720
Qy	1909 GTGATGATCATGACGAAGTATCTCTATACTGCACTTTTCAGTTCAGATGTTTACTGTTTC 1968
Db	721 GTGATGATCATGACGAAGTATCTCTATACTGCACTTTTCAGTTCAGATGTTTACTGTTTC 780
Qy	1969 CTACTCTGGCTCGACATCTTATTAAGAGCAGAAATGTTATCTCTGTCAATTACTGAAACTC 2028
Db	781 CTACTCTGGCTCGACATCTTATTAAGAGCAGAAATGTTATCTCTGTCAATTACTGAAACTC 840
Qy	2029 TGCTAGAGHTTTTACTGAGTACTTGGACAGGAACAATAAATTCAACTTCACGGGTTATA 2088
Db	841 TGCTAGAGHTTTTACTGAGTACTTGGACAGGAACAATAAATTCAACTTCACGGGTTATA 900

QY 3169 ATGAATCACTCAAGAGCTTCAATATGTACTTTTATCATTACTCCAAACCAGCATAGCA 3228
Db |||||||
QY 1981 ATGAATCACTCAAGAGCTTCAATATGTACTTTTATCATTACTCCAAACCAGCATAGCA 2040
Db |||||||
QY 3229 AGGCTGAACATATGCAGAGAAAGAGGAGAAACCAAGAAACCAAGAGATGAAGCATTTGCCGC 3288
Db |||||||
QY 2041 AGGCTGAACATATGCAGAGAAAGAGGAGAAACCAAGAAACCAAGAGATGAAGCATTTGCCGC 2100
Db |||||||
QY 3289 CACACCACTCTCGAATCTGCTGCTTTCAGCAAGATGATTAACCTTCTCAACTGTG 3348
Db |||||||
QY 2101 CACACCACTCTCGAATCTGCTGCTTTCAGCAAGATGATTAACCTTCTCAACTGTG 2160
Db |||||||
QY 3349 ATATCATGATGATCAATCTCAGGACCGTATTTGAGCGGCAATAGACACATTTCTAACT 3408
Db |||||||
QY 2161 ATATCATGATGATCAATCTCAGGACCGTATTTGAGCGGCAATAGACACATTTCTAACT 2220
Db |||||||
QY 3409 TGTGACCGAAGGATGCTCCAAATGGCTTTTTCATATTTCTGGCATTTGGGTTTACTAGAG 3468
Db |||||||
QY 2221 TGTGACCGAAGGATGCTCCAAATGGCTTTTTCATATTTCTGGCATTTGGGTTTACTAGAG 2280
Db |||||||
QY 3469 AGAAGCAACAGCTTCAAAAGCTCTGAAGAGATGAATTTGATTTTATCATTAAG 3528
Db |||||||
QY 2281 AGAAGCAACAGCTTCAAAAGCTCTGAAGAGATGAATTTGATTTTATCATTAAG 2340
Db |||||||
QY 3529 CTTCAAGATTTGGGAGTTTCAGCCATGATATACAAATGCTTTTGGAAAACTCAAAGGAA 3588
Db |||||||
QY 2341 CTTCAAGATTTGGGAGTTTCAGCCATGATATACAAATGCTTTTGGAAAACTCAAAGGAA 2400
Db |||||||
QY 3589 TTCCCCAGTTAGAGGCGCAGAGACATGATAAAGTGTGATATCTTCAGATGTTTGACACAG 3648
Db |||||||
QY 2401 TTCCCCAGTTAGAGGCGCAGAGACATGATAAAGTGTGATATCTTCAGATGTTTGACACAG 2460
Db |||||||
QY 3649 TGAAGCGATTAGAGAAAAATCTTTGTTAAATTTAGTACCAACCATATGAGATCGGAATCTA 3708
Db |||||||
QY 2461 TGAAGCGTTAAGAGAAAAATCTTTGTTAAATTTAGTACCAACCATATGAGATCGGAATCTA 2520
Db |||||||
QY 3709 TTAAGATGATGAGATTTACTCATGATTAAG 3738
Db |||||||
QY 2521 TTAAGATGATGAGATTTACTCATGATTAAG 2550
Db |||||||
RESULT 11
ACN39798
ID ACN39798 standard; cDNA; 3776 BP.
XX
AC ACN39798;
XX
DT 18-NOV-2004 (first entry)
XX
DE Tumour-associated antigenic target (TAT) cDNA DNAS25984, SEQ ID NO:4163.
XX
KW Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW tumour; diagnosis; cell proliferative disorder; breast cancer;
KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW central nervous system cancer; bladder cancer; pancreatic cancer;
KW cervical cancer; melanoma; leukaemia; hybridisation probe;
KW chromosome identification; chromosome mapping; gene mapping;
KW gene therapy; cytostatic; gene; ss.
XX
OS Homo sapiens.
XX
PN W02004030615-A2.
XX
PD 15-APR-2004.
XX
PP 29-SEP-2003; 2003WO-US028547.
XX
PR 02-OCT-2002; 2002US-0414971P.
XX
PA (GETH) GENENTECH INC.
XX
PI Wu TD, Zhang Z, Zhou Y;

DR WPI; 2004-347921/32.
XX P-PSDB; ABM81609.
PT New tumor-associated antigenic target polypeptides and nucleic acids,
PT useful in preparing a medicament for treating or detecting a
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT prostate cancer or tumor.
XX
PS Claim 1; SEQ ID NO 4163; 7273pp; English.
XX
CC The invention relates to human tumour-associated antigenic target (TAT)
CC polypeptides, and their related nucleic acids. The TAT polypeptides are
CC overexpressed in cancer tissues compared to normal tissues, and may thus
CC serve as effective targets for the diagnosis and treatment of cancer in
CC mammals. The invention also relates to nucleic acid and polypeptide
CC sequences at least 80% identical to the TAT nucleic acids and
CC polypeptides; expression vectors and host cells comprising a TAT nucleic
CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC TAT polypeptide; and methods and compositions for the treatment or
CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids, useful
CC antibodies, antagonists, binding molecules and compositions are
CC for diagnosing or treating a cell proliferative disorder associated with
CC increased TAT expression, particularly cancers such as breast cancer,
CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC cancer, pancreatic cancer, cervical cancer, cancers of the central
CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be
CC used as hybridisation probes, in chromosome and gene mapping, in
CC chromosome identification and in gene therapy. The present sequence
CC represents a TAT nucleic acid of the invention
XX
SQ Sequence 3776 BP; 1081 A; 703 C; 750 G; 1242 T; 0 U; 0 Other;
Query Match 27.2%; Score 1716.4; DB 13; Length 3776;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1717; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 4591 CACAAATTTATAGAAATTTGGATTGAAAGTGGCACTTGAAGTGGCACTTGAAGGATCTCGAGTCCCCA 4650
Db 1 CACAAATTTATAGAAATTTGGATTGAAAGTGGCACTTGAAGTGGCACTTGAAGGATCTCGAGTCCCCA 60
QY 4651 TGCTGACCTGGAGCACTCTGGGCTTTTCACTATCCAGGCAATTTGAAATCTATTGGGAGATG 4710
Db 61 TGCTGACCTGGAGCACTCTGGGCTTTTCACTATCCAGGCAATTTGAAATCTATTGGGAGATG 120
QY 4711 AAGGAAAACCTCTGTTTGGAGCACTTCAAAATAGGAGCATAATGGTCTGAAAGCATTA 4770
Db 121 AAGGAAAACCTCTGTTTGGAGCACTTCAAAATAGGAGCATAATGGTCTGAAAGCATTA 180
QY 4771 TGCAGTTTGCAGTTGTCACAGAGGATTTACCTGCTCAGTCTCTGATACAGAAACATCTGG 4830
Db 181 TGCAGTTTGCAGTTGTCACAGAGGATTTACCTGCTCAGTCTCTGATACAGAAACATCTGG 240
QY 4831 TTCTGCTTCTATCAGTTGTTCTTCTTAACATAAATCAGAAAGATACACATGCTTCTGT 4890
Db 241 TTCTGCTTCTATCAGTTGTTCTTCTTAACATAAATCAGAAAGATACACATGCTTCTGT 300
QY 4891 CTATAGATCTGTTTTCATGTTTGGTGGTGTGTGTAGCATTCCTATTCCTTGTATTGGG 4950
Db 301 CTATAGATCTGTTTTCATGTTTGGTGGTGTGTGTAGCATTCCTATTCCTTGTATTGGG 360
QY 4951 ATGACCTGTTGATCTGCAGGCTTCTTTCAGTTAGTTTCTTCTATTAACACCTTTTATCTCT 5010
Db 361 ATGACCTGTTGATCTGCAGGCTTCTTTCAGTTAGTTTCTTCTATTAACACCTTTTATCTCT 420
QY 5011 TCCATTTGATCACCATGCGCACACATGCTTTCAGATACTACTTACAGTAGACACAGCCCTAC 5070
Db 421 TCCATTTGATCACCATGCGCACACATGCTTTCAGATACTACTTACAGTAGACACAGCCCTAC 480
QY 5071 CCCTTGTCTCAGGTTCAAGAGACAGATGAGAGGCTCATTCGCGCATCTTCTTTTTCGAG 5130
Db 481 CCCTTGTCTCAGGTTCAAGAGACAGATGAGAGGCTCATTCGCGCATCTTCTTTTTCGAG 540

5131 AAATTTCTCAATATACAGTGGCTCCATTTGGGTGTGATATTCCTGGCTGGTATTTGTGGG 5190
Db AAATTTCTCAATATACAGTGGCTCCATTTGGGTGTGATATTCCTGGCTGGTATTTGTGGG 600
5191 TCTCACTGAAGAAATGGCACTACCCCTTATCTTCGCTGTGCTGCATTTGTTTCCACTATT 5250
Db TCTCACTGAAGAAATGGCACTACCCCTTATCTTCGCTGTGCTGCATTTGTTTCCACTATT 660
5251 TACTTGGGGTAACTCGGCTGAGGAATGTCATACCAATTTCTGCAGAGAGAGTACAGTG 5310
Db TACTTGGGGTAACTCGGCTGAGGAATGTCATACCAATTTCTGCAGAGAGAGTACAGTG 720
5311 CACTCTGTAGCTATCTATCTTACCTACAAATTTCTCTCTCTCTCCAGGAATTTGGG 5370
Db CACTCTGTAGCTATCTATCTTACCTACAAATTTCTCTCTCTCTCCAGGAATTTGGG 780
5371 ATACTGTAAAGCCCTTGTCTCAGAGGTGTGTGCAGATCTCTGCTTACTAACTGTTGA 5430
Db ATACTGTAAAGCCCTTGTCTCAGAGGTGTGTGCAGATCTCTGCTTACTAACTGTTGA 840
5431 AGCAAAAAACACCGTGTGAGTACCTTAGAAAAAGAAATAGTTTGTATAGAGCTTCTG 5490
Db AGCAAAAAACACCGTGTGAGTACCTTAGAAAAAGAAATAGTTTGTATAGAGCTTCTG 900
5491 ATGACTATAGCTGCTCTCTGATCAAGCTTCTCATTTCAAGTGCACCGTCTGCAGATG 5550
Db ATGACTATAGCTGCTCTCTGATCAAGCTTCTCATTTCAAGTGCACCGTCTGCAGATG 960
5551 ATGACGAAAGCATCTCTGCT 5610
Db ATGACGAAAGCATCTCTGCT 1020
5611 TTTGCTGCAGGAAATTTGAAACGGGGAAGGTTGGAGCTTGTCAATTTTTCACGCACTTC 5670
Db TTTGCTGCAGGAAATTTGAAACGGGGAAGGTTGGAGCTTGTCAATTTTTCACGCACTTC 1080
5671 ACTGTGGAGCGGAGTCTGCATTTTCTTAAATCAGAAATGCCGAGTGTCTGCTGTTG 5730
Db ACTGTGGAGCGGAGTCTGCATTTTCTTAAATCAGAAATGCCGAGTGTCTGCTGTTG 1140
5731 AAGGTAAGCCAGAGGCTGTGCTATCCAGCTTCTTACTTGGATGAATATGAGAAACAG 5790
Db AAGGTAAGCCAGAGGCTGTGCTATCCAGCTTCTTACTTGGATGAATATGAGAAACAG 1200
5791 ACCCTGGCTCAAGAGGGGCAACCCCTTCAATTTATCTCGTGGCGGTATCGGAGCTCC 5850
Db ACCCTGGCTCAAGAGGGGCAACCCCTTCAATTTATCTCGTGGCGGTATCGGAGCTCC 1260
5851 ATTTGGCTGCGCAACACACTGCAATATAGAGAGATTTGCTAGGAGCCAGAGACTAATC 5910
Db ATTTGGCTGCGCAACACACTGCAATATAGAGAGATTTGCTAGGAGCCAGAGACTAATC 1320
5911 AGATGTTATTTGATTTCAACTGGCAGTGTACTGTGAGCTTCAACTCTGCTCAAGCAATC 5970
Db AGATGTTATTTGATTTCAACTGGCAGTGTACTGTGAGCTTCAACTCTGCTCAAGCAATC 1380
5971 ACAATGACGACAGTAGTAAGGCTGATTTCAAAATTTATGGAATCTTCTGAGGGCTGGG 6030
Db ACAATGACGACAGTAGTAAGGCTGATTTCAAAATTTATGGAATCTTCTGAGGGCTGGG 1440
6031 AAGATTTGGAGGGGCTTTTGTCTCAATCTCAGGTTTCACTTACATCAATAAATTTTCT 6090
Db AAGATTTGGAGGGGCTTTTGTCTCAATCTCAGGTTTCACTTACATCAATAAATTTTCT 1500
6091 TAATGGAGTATGCTTTCAATTAGCAACATATGCTTACAGGAAAAAGACATAGATC 6150
Db TAATGGAGTATGCTTTCAATTAGCAACATATGCTTACAGGAAAAAGACATAGATC 1560
6151 AATCTGTTTATGCTAGTATTTCCAGGAATTTATTTCCCTTCTCATATTTGTCTCATTT 6210
Db AATCTGTTTATGCTAGTATTTCCAGGAATTTATTTCCCTTCTCATATTTGTCTCATTT 1620
6211 CATTTTATTCATCCACTTGGTAGATGAAGTCACGTCAACAGTTGTAGACATTTTATGT 6270

Db 1621 CATTTTATTCATCCACTTGGTAGATGAAGTCACGTCAACAGTTGTAGACATTTTATAT 1680
Qy 6271 GTTGGTTAACTCTCTCTGCAATTTTGTATTTGTGTTT 6308
Db 1681 GTTGGTTAACTCTCTCTGCAATTTTGTATTTGTGTTT 1718

RESULT 12

ADRO7317
ID ADRO7317 standard; cDNA; 3598 BP.
XX AC ADRO7317;
XX DT 04-NOV-2004 (first entry)
XX Full length human cDNA useful for treating neurological disease Seq 823.
DE gene; ss; human; oligo-capping method; diagnostic marker; gene therapy;
KW osteoporosis; neurological disease; Alzheimer's disease;
KW Parkinson's disease; dementia; short memory; cancer;
KW sense or motor function; emotional reaction; fear response; panic;
KW osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;
KW tranquiliser.
XX Homo sapiens.
OS
XX
PN EP1447413-A2.
PD 18-AUG-2004.
XX
PF 12-FEB-2004; 2004EP-00003145.
XX
PR 14-FEB-2003; 2003JP-00102207.
PR 09-MAY-2003; 2003JP-00131452.
XX
XX (REAS-) RES ASSOC BIOTECHNOLOGY.
PI Isogai T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T;
PI Wakamatsu A, Ishii S, Nagai K, Irle R;
XX WPI; 2004-583265/57.
DR P-PSDB; ADRO9273.
XX
PT New 1995 cDNA, useful for treating osteoporosis, neurological diseases,
PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX
PS Claim 1; SEQ ID NO 823; 2686pp; English.
XX

This invention relates to novel, isolated full length human cDNA molecules and the encoded proteins thereof. Specifically, it refers to cDNA clones obtained by an oligo-capping method, where none of these clones are identical to any known human mRNAs. The present invention describes an immunoassay to identify agonists and antagonists, as well as antibodies, antisense molecules and siRNAs that can all be used to bind to and modulate expression of the cDNA molecules. As such, these molecules are useful for diagnostic markers or therapeutic targets for the various diseases or morbid states. In particular, they are useful in gene therapy for treating osteoporosis, neurological disease, Alzheimer's disease, Parkinson's disease, dementia, short memory and various cancers, as well as for maintaining equilibrium of sense or motor function, and for treating emotional reaction, fear response and panic. Accordingly, they exhibit osteopathic, neuroprotective, nootropic, antiparkinsonian, cytostatic and tranquiliser activities. This polynucleotide is a full length human cDNA sequence of the invention. NOTE: This sequence is not given in the sequence listing of the specification but can be obtained on CD-ROM from the European Patent Office, Vienna Sub-office.

XX SQ Sequence 3598 BP; 1027 A; 669 C; 709 G; 1193 T; 0 U; 0 Other;

Query Match 24.2%; Score 1529.2; DB 13; Length 3598;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1531; Conservative 0; Mismatches 3; Indels 0; Gaps 0;


```
QY 4775 GTTTCAGTTGCACAGAGATTACCTGCTCAGGTCTGTATACAGAAACATCTGTTTCG 4834
Db 1 GTTTCAGTTGCACAGAGATTACCTGCTCAGGTCTGTATACAGAAACATCTGTTTCG 60
QY 4835 TCTTCTATCAGTTGTTCTTCTTAAACATAAAATCAGAAGATACACCATGCTTCTGTCTAT 4894
Db 61 TCTTCTATCAGTTGTTCTTCTTAAACATAAAATCAGAAGATACACCATGCTTCTGTCTAT 120
QY 4895 AGATCTGTTTCAATGTTTGGTGGTGTGTGTAGCATTCATCCATCCTTGTATGGGATGA 4954
Db 121 AGATCTGTTTCAATGTTTGGTGGTGTGTGTAGCATTCATCCATCCTTGTATGGGATGA 180
QY 4955 CCTCTGTACCTGACAGCTTCTTCAAGTGTAGTTCTTCTTATACACCATCTTATCTTCCA 5014
Db 181 CCTCTGTACCTGACAGCTTCTTCAAGTGTAGTTCTTCTTATACACCATCTTATCTTCCA 240
QY 5015 TTTGTATACCATGGCACATGCTTTCCAGTACTTACAGTAGACACAGGCTTACCCCT 5074
Db 241 TTTGTATACCATGGCACATGCTTTCCAGTACTTACAGTAGACACAGGCTTACCCCT 300
QY 5075 TGCTCAGGTTCAAGAAGACAGTGAAGGCTCATTTCCGATCTTCTTCTTTGCGAAT 5134
Db 301 TGCTCAGGTTCAAGAAGACAGTGAAGGCTCATTTCCGATCTTCTTCTTTGCGAAT 360
QY 5135 TTCTCAATATACAGTGGCTCCATTTGGGTGTGATATCTGCTGGTATTTGTGGGTCTC 5194
Db 361 TTCTCAATATACAGTGGCTCCATTTGGGTGTGATATCTGCTGGTATTTGTGGGTCTC 420
QY 5195 ACTGAAGATGGCATCACCCCTTATCTTCTGCTGTCTGATTTCTTCTTCTTCTTACT 5254
Db 421 ACTGAAGATGGCATCACCCCTTATCTTCTGCTGTCTGATTTCTTCTTCTTCTTACT 480
QY 5255 TGGGTATATCTGCTGTAGGAACTGCATACCAATTTCTGAGAGAGAGTACAGTGCAT 5314
Db 481 TGGGTATATCTGCTGTAGGAACTGCATACCAATTTCTGAGAGAGAGTACAGTGCAT 540
QY 5315 CTGTAGCTATCTATCTTACCTACCAATTTGTTCTGCTCTTCCAGGAATATTTGGATAC 5374
Db 541 CTGTAGCTATCTATCTTACCTACCAATTTGTTCTGCTCTTCCAGGAATATTTGGATAC 600
QY 5375 TGTAAGGCCCTTGCTCCAGAGGTGTGTGCAGATCCTGCTTACTTAAACTTGTTTGAAGCA 5434
Db 601 TGTAAGGCCCTTGCTCCAGAGGTGTGTGCAGATCCTGCTTACTTAAACTTGTTTGAAGCA 660
QY 5435 AAAAAACACCGTGTGAGTACCTTAGAAAAAAGAAATAGTTGTAGAGCTTCTCTGATGA 5494
Db 661 AAAAAACACCGTGTGAGTACCTTAGAAAAAAGAAATAGTTGTAGAGCTTCTCTGATGA 720
QY 5495 CTATAGCTGCTCTGGAATCAAGCTTCTCATTTTCAGGTGCCACGGTCTGCAGATGATGA 5554
Db 721 CTATAGCTGCTCTGGAATCAAGCTTCTCATTTTCAGGTGCCACGGTCTGCAGATGATGA 780
QY 5555 GCGAAGAGCATCTGCTCTGCTCTGCTTCTGTTGGGCTATATCTTCTCAGAAATTTG 5614
Db 781 GCGAAGAGCATCTGCTCTGCTCTGCTTCTGTTGGGCTATATCTTCTCAGAAATTTG 840
QY 5615 CTGCCAGGAATTTGTGAACGGGGAAGAGTTGGAGCTTGCAATTTTTCAGCATCTTCACTG 5674
Db 841 CTGCCAGGAATTTGTGAACGGGGAAGAGTTGGAGCTTGCAATTTTTCAGCATCTTCACTG 900
QY 5675 TGGAGCCGGAGTCTGCAATTTTCTTAAATCAGAGAAATCCGAGTGGTCTGTTGAAGG 5734
Db 901 TGGAGCCGGAGTCTGCAATTTTCTTAAATCAGAGAAATCCGAGTGGTCTGTTGAAGG 960
QY 5735 TAAAGCCAGAGGCTGTGCTATCTTCCAGCTCTTACTTGGATGAATATGAGAAACAGACCC 5794
Db 961 TAAAGCCAGAGGCTGTGCTATCTTCCAGCTCTTACTTGGATGAATATGAGAAACAGACCC 1020
QY 5795 TGGCTGTGAAGAGGGCAACCCCTTCAATTTATCTGTGAGCGGTATCGAAGCTCCATTT 5854
Db 1021 TGGCTGTGAAGAGGGCAACCCCTTCAATTTATCTGTGAGCGGTATCGAAGCTCCATTT 1080
```

```
QY 5855 GGTCTGCGAACAAACATCTGCATTTATAGAAGAGATTGTCTAGGAGCCAGAGACTAATCAGAT 5914
Db 1081 GGTCTGCGAACAAACATCTGCATTTATAGAAGAGATTGTCTAGGAGCCAGAGACTAATCAGAT 1140
QY 5915 GTTATTTTGGATTCAACTGGCGAGTTACTGTGAGCTTCCAACTCTGCTCAAGCAATCAAA 5974
Db 1141 GTTATTTTGGATTCAACTGGCGAGTTACTGTGAGCTTCCAACTCTGCTCAAGCAATCAAA 1200
QY 5975 ATGACGACAGTAGTAAAGGCTGATTCAAAATTTATGGAATACTTTCTGAGGCTCGGAAAG 6034
Db 1201 ATGACGACAGTAGTAAAGGCTGATTCAAAATTTATGGAATACTTTCTGAGGCTCGGAAAG 1260
QY 6035 TATTGAGGGTCTTTTGTCTCCATGTCCAGGTTCACTTACATCAATAAATAATTTCTTAAT 6094
Db 1261 TATTGAGGGTCTTTTGTCTCCATGTCCAGGTTCACTTACATCAATAAATAATTTCTTAAT 1320
QY 6095 GGAGTATTTGCTTTCAATTAGCAAAACATATGCTTCCAGGAAAGGAAAGCATAGATCAATC 6154
Db 1321 GGAGTATTTGCTTTCAATTAGCAAAACATATGCTTCCAGGAAAGGAAAGCATAGATCAATC 1380
QY 6155 TGTATTTATGTGCTAGTATTTCCAGGAATTTATTTCCCTTCAATAATTTGTCTCATTTCAAT 6214
Db 1381 TGTATTTATGTGCTAGTATTTCCAGGAATTTATTTCCCTTCAATAATTTGTCTCATTTCAAT 1440
QY 6215 TTATTTTCAATCCATCTGCTAGTATGAGTCAAGTCAAGTCAAGGTTGTAGACATTTTATGTGTTG 6274
Db 1441 TTATTTTCAATCCATCTGCTAGTATGAGTCAAGTCAAGTCAAGGTTGTAGACATTTTATGTGTTG 1500
QY 6275 GTTAACTCTTCTGCAATTTTGTATTTGGTGTGTTT 6308
Db 1501 GTTAACTCTTCTGCAATTTTGTATTTGGTGTGTTT 1534

RESULT 13
ADJ95443
ID ADJ95443 standard; cdna; 6300 BP.
XX
AC ADJ95443;
XX
DT 03-JUN-2004 (first entry)
XX
DE Human cdna encoding Ubiquitin ligase E3alpha II.
XX
KW Human; ss; gene; ubiquitin ligase; E3alpha II;
KW ubiquitin-proteasome pathway; gene therapy; vaccine; muscular atrophy;
KW cachexia; catabolic disorders; cancer cachexia; renal cachexia;
KW inflammatory cachexia; muscle wasting disorder; metabolic acidosis;
KW uremia; burn; hyperthyroidism; Cushing's syndrome; fasting;
KW denervation atrophy; diabetes mellitus; sepsis; AIDS wasting syndrome.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 295..5562
FT FT /*tag= a
FT FT /product= "Ubiquitin ligase E3alpha II"
XX
PN US6706505-B1.
XX
PD 16-MAR-2004.
XX
PP 28-NOV-2000; 2000US-00724126.
XX
PR 08-MAR-2000; 2000US-0187911P.
XX
PA (AMGE-) AMGEN INC.
XX
PI Han H, Kwak K;
XX
XX WPI; 2004-236723/22.
DR P-PSDB; ADJ95444.
XX
PT New nucleic acid molecule, useful for preparing a composition for
```

PT diagnosing, treating or preventing diseases associated with human
 PT E3aproximatelya polypeptide, e.g., muscle atrophy.
 XX
 PS
 XX Example 2; SEQ ID NO 3; 104pp; English.
 CC The invention relates to a new isolated nucleic acid molecule appearing
 CC as ADJ95441 (or its complement) encoding a ubiquitin ligase E3alpha I
 CC protein appearing as ADJ95442. Also included are a vector comprising the
 CC nucleic acid, a host cell comprising the vector, a process of producing a
 CC E3alpha I ubiquitin ligase polypeptide, a composition comprising the
 CC nucleic acid molecule, a reagent comprising a detectably labelled
 CC nucleotide, and a method for determining the presence of a human E3alpha
 CC I ubiquitin ligase nucleic acid in a biological sample. The nucleic acid
 CC molecule is useful for preparing a composition for diagnosing, treating
 CC or preventing diseases associated with human E3alpha I polypeptide, e.g.
 CC muscle atrophy, cachexia, catabolic disorders, cancer cachexia, renal
 CC cachexia, inflammatory cachexia, muscle wasting disorders associated with
 CC metabolic acidosis, uremia, burns, hyperthyroidism, Cushing's syndrome,
 CC fasting, denervation atrophy, diabetes mellitus, sepsis and AIDS wasting
 CC syndrome. The present sequence encodes human E3alpha II.
 XX
 SQ Sequence 6300 BP; 1842 A; 1235 C; 1478 G; 1745 T; 0 U; 0 Other;

Query Match 21.9%; Score 1381.8; DB 12; Length 6300;
 Best Local Similarity 56.5%; Pred. No. 0;
 Matches 2947; Conservative 0; Mismatches 2167; Indels 105; Gaps 16;
 780 TGGGATCAGCAAGTGTGATTTTATATCTGCTTTCTGTCATCATTTGGCCAAATTTGGTGCCA 839
 376 TGGCTGCAAGCACTGACCTCACTAGAGAGTGTACCGATTTAGCCCACTATGTACCC 435
 840 GAAATTTACTTTGCTGAAATGACCCAGACTTTGGAAAAGCAGGAGGAAGTGTACAAATG 899
 436 AAAATCTACTGCGGGGTCCCAACCTTTTCCACAGAAAGACATGCTGGCACAGCAT 495
 900 TCAATTTACTCCACTGGAATGGTACTTTATTTGGAGAAGATCCAGATATTTGCTTAGAG 959
 496 GTTTTGTGGGACCAATGGATGGTACCTTTGTGGTGAAGATCTGCAATTTGGATTTCCA 555
 960 AAATTTGAAGCAC- --AGTGGAGCATTTTCAGCTTTTGGGAGGGTTTCAAAGTGGAGAG 1016
 556 AAATTTGAGCAAGCAACAACTTCTCATCTTTGTGGTGGTGTGTTTAAAGTAGAGAG 615
 1017 ACAACCTATTTCTGAGGAGTGTGCAATTTGATCCAAATGATGTACTCTGTATGGAAGTGC 1076
 616 CCTACATATTTCTGAGAGACTGTGCAATTTGATCCAACTTTGTTTGTGCAATGAGTGC 675
 1077 TTCCAGGACAGTGTTCATAAAATCATCTGTTTCAAGATGCATCTTCTACTGGAGAGGG 1136
 676 TTTTGGGAAGTATTTACAGAGATCATCGATATAGGATGACAAATCAGGAGGTGGAGGT 735
 1137 TTCTGTGACTGTGGAGACACAGGAGCATGGAAATCTGGCCCTTTTGTGTAATCATCAA 1196
 736 TTCTGTGACTGTGGTGTACTGAGACCTGGAAAGAGGGTCTTACTGTCAAAAACATGAA 795
 1197 CTTGGAAGAGCAGGTACT- --ATAAAGAGATTTCAAGTGTGCTGCTGTAATGAAGAGGTA 1253
 796 CTTAACACCTCTGAAATTTGAGGAAGAGAGGATCTCTTGTTCATTTATCAGAAGATGTG 855
 1254 ATTGTCCAGCAGGAAATATTTCTTCAGTGATGATAAATATGTCTGTAGAAATGATCTATA 1313
 856 ATAGCAAGAACTTATAACATTTTGTCTATTTACGTTTTCGTTATGCAAGTATGAAATATTAACC 915
 1314 TGGGAAGAGGAAAGAAAGTGCCTCTGCACTCCAGATGAAGGAGGAAATGAAGATAC 1373
 916 TGGGAAAGAAAGTGAATTTGCCAGAGATTTAGAGATGGTAGAGAGTACACCTTAC 975
 1374 TATTGTGCTTTTCAATGATGAACACCAATTCATATGACCAAGTGTATATACAGCCCTACAA 1433
 976 TATTGCACTGCTTTTAAATGATGAGGTTTACACCTATATGAACAGTATTTTATCTCTTAC 1035
 1434 AGAGCTTTGACTGTGAGCTGCGAGGCGCCAGTTGTCATACACTGCGCAATTTGACAAAGAG 1493

Db 1036 AAAGCTGTTAACTGTACACAAAAGAACGCTATTGTTTGGCACTACAGTAGATCGAGAT 1095
 QY 1494 GGTCTGGGCTGTTTAAAGCGGGAGCTTTATGCTGCTTCCAGGAGCAAGAGGAGATATA 1553
 Db 1096 GGGCTAGTCTGTTTCGATATGGAGATTTTCAGTATTTGTGAGCAAGCAAAATCAGTAAT 1155
 QY 1554 AAGAGTCAATTCAGAAAATGCTCTCAACATCCACTTCATGTAGAGATTTATACACTCAGAG 1613
 Db 1156 GTGGAATATACCAAGTAGACAGAC- --AAAGCCACTCAAAGTTCAGATTTATGCTTCTGCT 1212
 QY 1614 ATTATGGCTCATCAGAAATTTGCTTGGCTTCTGCTGCTGATCAACAAAATATATGAGC 1673
 Db 1213 ATTGTGCAATCAGAAATTTGCTTGGTGAACCTTTGCTGCTGGGAGATTTATTTGA 1272
 QY 1674 TATTCAAGTGAATTTTAGGCAGATCTTTTGGCAAGCATGCTTTAGAGAAAGAACCTGACTCG 1733
 Db 1273 TATTCAAGTGCCTTTGCGCGGATTTTATGTCAAGTTGGTTTACAAGAGGGCCAGATGCT 1332
 QY 1734 GAGATCCCTGCTCATTAAGCAGGTTTATGCTTTGGGATGCAAGCTTTTATAAGGTGCC 1793
 Db 1333 GAAAACCTCTTCTCTAGTGGACAGACTGATGCTTAGTATCCAAATTTATGGAAGGTGCT 1392
 QY 1794 CGTAAGATCTTTCATGAATTCATCTTCAAGCAGTTTATTTTATGGAGATGGAATACAAAAA 1853
 Db 1393 AGGAGTGTATATCATCAGTTGTTTATGAGCAGTCTGCTTATGGATTTGAAATACAGAAA 1452
 QY 1854 CTCTTGTCTATGGAATTTTGAAGTATTTATAAACAACCTGCAAGAAATATATCAAGTAT 1913
 Db 1453 CTATTTGCTGTTTCGATTTGCAAAAAATTTACCAGCAGTTTGCAGAGAGATTTTATGGAGAT 1512
 QY 1914 GATCATGACAGAGTATCTTATTAACCTGCACTTTCAAGTTCAGATGTTTACTGTTCTTACT 1973
 Db 1513 GATCAGCAGCAGAGTGTGCTGCTCTATCTGCTCCAGTCTTCCACGCACTTACT 1572
 QY 1974 CTGCTGACATCTTATTTGAAGCAGATGTTATCTCTGTCATTTACTGAACTCTGCTA 2033
 Db 1573 CTGCTGCAATGCTCATCAGAGAAATCTTGAATGAGCATTTATCAATTAAGCTTTTATG 1632
 QY 2034 GAAGTTTT- --ACCTGAGTACTTGGACAGGAACAAATAAATTCACCTTCCAGGCTTAT 2087
 Db 1633 GATCATTTGAGACATCGAGATGCCAGGGCAGATTTCAAGTTTGAACGATACACTGCTTTA 1692
 QY 2088 AGCCAGGACAAATTTGGGAAGAGTATATGCAATATATGTGACCTTAAAGATATCTGATC 2147
 Db 1693 CAAGCTTCAAATTTTAGGAGAGTACAGAGCCTTATTTTAGATCTCAAGATATGTTAAAT 1752
 QY 2148 AGCAAAACCAATATTTGACAGAAAGATTAAGATGCAATGCTTCTTGAAGGTTTTCGATCT 2207
 Db 1753 AGCAAAACCAATGCTGATGAGCTGAGGAGGAGGTTTCTAGAAAGGTTTGTATGCC 1812
 QY 2208 TTTTGAAGATTTCTTACCTGTATGCGGGAATGGAAGAAATCCGAAGACAGGTTGGGCAA 2267
 Db 1813 TTTTGGATTTACTTAAATTTATGAGGAGATGATCCATTTACAGTCAAGTAGGACAA 1872
 QY 2268 CACATTAAGTGGATCTTGAATTTGGAGGCTGCCATTTGCTATACAGATGCAATTTGAAGAT 2327
 Db 1873 CATATTGAATGGAACCAAGAGTGGGAGAGCCTTTCACACTACAAATGAAATTAACAT 1932
 QY 2328 ATTTTACTCATGTTTCCAGAGTGGTGTGCTGCTGATGAGAGACTTCTTACTTGTGGCTTAT 2387
 Db 1933 GTCAATTTCAATGATGAGGAGTGGTGTGCTTCAAGTGAAGAGGTTTAAATCGAAGCTTAC 1992
 QY 2388 AAAGAAATGTCACAAAGCTGTGATGAGGTCAGTACCAGTTTTCATATCTAGTAGCAAGACA 2447
 Db 1993 AAGAAATGTCGCTGCTGCTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2052
 QY 2448 GT- --AGTACAACTGCTGTGGAATAGTTTGGAAACAAAGTCTCTACAGAGTATCTGAGAT 2504
 Db 2053 ATCACACTAAGCAATTTTGTGGACATTTAGTGGAACTATCAGATACTGTTGTTTCCCAAGAA 2112
 QY 2505 CTTGTAGACATACATCTGCGCACTCTCTAGGAGCCCTTGTGCTGCTTCTCATGTAGGTTTAAAGC 2564
 Db 2113 AAAGTTAGCAATTCACCTCCCAAGTTTCTGCTTACTTGCAGGTTTACATGATTTATTAAGC 2172

Db 4312 ATCCAAAGCATAGAAAGAAATTTGAGTGAATGAAGATAAACCAATTTGTTGGCTCTTACCT 4371
QY 4740 AATAGCAGCATAAATGGTCTGAAGCAATTAATGCAAGTTTGGCAGAGGAAATACC 4799
Db 4372 TGCAGACTGGATGACTGTCTTAGGTCAATGACAGATTTGGCCGACGACACTGGACAGTG 4431
QY 4800 TGCTCTCAGTCTGTATGACAGAAACATCTGGTGTGCTCTATCATAGTGTCTCTCTTAAC 4859
Db 4432 GCATCAGTTTTCAGTGGTGCAGAGCAATTTTGTAAACCTTTTGGCATCACTGGTGCCTAAT 4491
QY 4860 ATAAATCAGAAAGATACACATGCTCTGTCTATAGATCTGTTTCATGTTTGGTGGCT 4919
Db 4492 GACAGCATGAGAACTTCCATGCAATATAGATATATGACATGTTTCATTTATTTGGTGGC 4551
QY 4920 GCTGTGTTAGCAATCCCAATCTGTATATGGGATGACCCCTGTTGATCTGCAGCCTTCTTCA 4979
Db 4552 TTGGTGTGTCATTTCTGCTGGTGCAGTG-----TCAGGATTTTTCAGGG 4596
QY 4980 GTTAGTCTCTCTATTAACACCTTTATCTCTTCCATTTGATCACAATGGCACAATGCTT 5039
Db 4597 ATCAGCTTTGGCACTGGAGACCTTCCATTTTCCATCTGTTTACTATATGGCACAATCATA 4656
QY 5040 CAGATACTACTACAGTACAGACAGCC---TACCCTTGTCTCAGTTCACAGACAGT 5096
Db 4657 CAGATCTTACTTACTCTGTATCAGAGAGAAATGGCATGGATCAAGAAATCCCTCTGT 4716
QY 5097 GAAGAGGCTCATTCGCGACTCTTCTTTTTCAGAAATTTCTCAATATACAGTGGCTCC 5156
Db 4717 GAAGAAGATCAGCAGTCTTCTGTTGTATTAACACTTCCAGCATATACGGGAGTGCC 4776
QY 5157 ATTGGGTGTGATATTCCTGGCTGGTATTTGTGGGTCTCACTGAAGATGGCAATCCCT 5216
Db 4777 TTGAAGAAATACCATCCGCTGGCATCTGTGGAGGAGTGTCAAGCTGGAATCATGCT 4836
QY 5217 TATCTTGGCTGTGCTGCTATGTTTTCACATTTTACTTGGGTAACTCCGCTCAGNA 5276
Db 4837 TTCTGGAAGTGTCTGCTTTATTTTTCATTAATTAATGAGTTCCTTCCCAACCCGAC 4896
QY 5277 CTGCATACCAATCTCTGAGAGGAGATACAGTGCATCTGTAGCTATCTATCTTTACT 5336
Db 4897 ATTCA---AGTTCCTGGAACAGCCATTTTGAACATTTATGTAGCTATCTTCCCTACCA 4953
QY 5337 ACAATTTGTTCTCTCTTCCAGGAATATGGGATCTGTAAAGCCCTTGTCTCAGAGG 5396
Db 4954 AACCACTCATTTGCTTTTCAAGAAATAGTGAATATGAATTCATGATGAAAGT 5013
QY 5397 TGGTGTGAGTCTGCTTACTTAACTGTTTGAAGCAAAACACCGTGTGAGTAC 5456
Db 5014 TGGTCCGCTAACAGTGAAGTTAAAGATATCTAGAAGGTGAAGAGATGCTATAGATAT 5073
QY 5457 CTTAGAAAAGAAATAGTTTGTATAGAGCTTCTGTATGATCTATAGCTGCTCTGAATCAA 5516
Db 5074 CCAAGAGATCTAACAAATTAATAACCTTCCAGAGATTACAGAGCTCTAATTAATCAA 5133
QY 5517 GCTTCTCTATTCAGGTGCCCGCTCTGCAGATGATGACGGAAGCATCTGTCTCTGTC 5576
Db 5134 GCATCCAAATTTCTGTCGCCGAAATCAGTGTGTATAGAGCAGAGCCCAACTCTGTGC 5193
QY 5577 CTTTCTGTGGGCTATATATGTTCTCAGAAATTTGTCGACGAAATTTGTGAACGGG 5636
Db 5194 CTTGTGTGGGATCTCTGCTGTGCTCCAGAGTTACTGCTGCCAGACTGAATGGAAGGG 5253
QY 5637 GAAGAGTTGGAGCTTGCATTTTTCAGGCATTTCTGTTGGAGCGGAGTCTGCAATTTTC 5696
Db 5254 GAGGATGTAGAGGCTCTCAGACTCACACTACTCTGTGGCTCTGAGTGGGCACTTTC 5313
QY 5697 CTAATAATCAGAAATGCGAGTGTCTGTTGAAGGTAAAGCAGAGGCTGTGCTCTAT 5756
Db 5314 CTGAGAGTACGGGAATGTGAGTGTCTATTTTGTAGCTGGCAAAACCAAGGCTGTTTAT 5373
QY 5757 CAGCTCTTACTTGGATGAATATGAGAAACAGACCTTGGCTGTGAAGGGGCAACCC 5816

Db 5374 TCTCTCTCTTACCTTGTGACTATGGGAGAGACCGACGAGGACTCAGACGGGGAATCCT 5433
QY 5817 CTTCAATTTATCTGTGAGCGGTATCGGAAGCTTCCATTTTGGTGTGGCAACAACACTGCATT 5876
Db 5434 TTACATTTATGCAAGAGCGATTCAGAAAGATTCAGAAAGCTCTGGCACCACACAGTGTG 5493
QY 5877 ATAGNAGAGATGCTAGGAGCGCAGAGACTAATCAGATGTTATTTGGATTCAACTGGCA 5935
Db 5494 ACAGAGGAATTTGGACATGTCACAGGAAGCCAAATCAGACTGGTTGGCAATTCAGTGGCA 5552

RESULT 14
ADS86865
ID ADS86865 standard; cDNA; 6300 BP.
XX
AC ADS86865;
XX
DT 16-DEC-2004 (first entry)
XX
DE cDNA encoding human E3alpha ubiquitin ligase, huE3alphaII.
XX
KW Human; E3alpha ubiquitin ligase; huE3alphaII;
KW ubiquitin-proteasome pathway; chromosome 15 q; rapid muscle wasting;
KW fasting; metabolic acidosis; muscle degeneration; kidney failure;
KW renal cachexia; uremia; diabetes mellitus; sepsis; AIDS wasting syndrome;
KW cancer cachexia; negative nitrogen balance; burn; Cushing's syndrome;
KW inflammatory cachexia; hyperthyroidism; denervation atrophy;
KW protein/tissue wasting; energy-protein malnutrition; muscle atrophy;
KW gene therapy; ss; gene.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 5'UTR 1..294
FT CDS /*tag= a
FT 295..5562
FT /*tag= b
FT /*product= "huE3alphaII"
FT 5563..6300
FT /*tag= c
XX
US2004185037-A1.
XX
PD 23-SEP-2004.
XX
PF 15-JAN-2004; 2004US-00758672.
XX
PR 08-MAR-2000; 2000US-0187911P.
XX 28-NOV-2000; 2000US-00724126.
XX (HANH/) HAN H.
XX (KWAK/) KWAK K.
XX Han H, Kwak K;
XX WPI; 2004-707854/69.
XX P-PSDB; ADS86866.
XX Novel isolated human E3alpha ubiquitin ligase nucleic acid molecule
XX useful for treating and/or preventing renal cachexia or inflammatory
XX cachexia.
XX
PS Claim 1; SEQ ID NO 3; 115pp; English.
XX
CC The present invention relates to new orthologues of human E3alpha
CC ubiquitin ligase, huE3alphaI and huE3alphaII. Most intracellular proteins
CC are degraded through the ubiquitin-proteasome pathway. Proteins are
CC marked for proteasomal degradation by conjugation of ubiquitin to the
CC protein. Conjugation of the ubiquitin molecule involves the activation by
CC E1 enzyme, subsequent transfer to E2 enzyme, which serves as a carrier,
CC and then interacts with a specific E3 ligase family member. E3 ligase
CC binds to proteins targeted for degradation and catalyses the transfer of
CC ubiquitin from the E2 carrier enzyme to the target protein. E3 ligase

QY 2247 ATCCGAAGACAGTTGGGCAACACATTTGAAGTGGATCCTGATTTGGGAGCTGCCATTGCT 2306
Db 121 ATCCGAAGACAGTTGGGCAACACATTTGAAGTGGATCCTGATTTGGGAGCTGCCATTGCT 180
QY 2307 ATACAGATCAATTTGAAGAAATATTTTACTCATGTTCCAAAGAGTGGTGTCTTTGTCATGAA 2366
Db 181 ATACAGATCAATTTGAAGAAATATTTTACTCATGTTCCAAAGAGTGGTGTCTTTGTCATGAA 240
QY 2367 GAACTCTTACTTGTGGCTTTATAAGAAATGTCTCAAAAGCTGTGATGAGGTGCAGTACCAGT 2426
Db 241 GAACTCTTACTTGTGGCTTTATAAGAAATGTCTCAAAAGCTGTGATGAGGTGCAGTACCAGT 300
QY 2427 TTCTATATCTAGTAGCAGCAGTACTACATCGTGTGGACATAGTTTGGAAACAAAGTCC 2486
Db 301 TTCTATATCTAGTAGCAGCAGTACTAGTACAAATCGTGTGGACATAGTTTGGAAACAAAGTCC 360
QY 2487 TACAGAGTATCTGAGGATCTTGTAAAGCAGTCTTGTCAAGCTCTCTAGGACCCCTGCTGCT 2546
Db 361 TACAGAGTATCTGAGGATCTTGTAAAGCAGTCTTGTCAAGCTCTCTAGGACCCCTGCTGCT 420
QY 2547 CTTCAAGTACGTTTAAAGCAGGCTGGTGTCTTTCAAGCTGTCATGAATTTGTGTCTTTT 2606
Db 421 CTTCAAGTACGTTTAAAGCAGGCTGGTGTCTTTCAAGCTGTCATGAATTTGTGTCTTTT 480
QY 2607 GAGGACTTTCAGTAGAGTACTAGTGGAAATCTTTACGTTGTCTGGTGTGGTGGCC 2666
Db 481 GAGGACTTTCAGTAGAGTACTAGTGGAAATCTTTACGTTGTCTGGTGTGGTGGCC 540
QY 2667 CAGGTTGTGTGAGATGTGGGAGAGAAATGACCTGTCTTTATTTAGCCAGGTGTTTTAT 2726
Db 541 CAGGTTGTGTGAGATGTGGGAGAGAAATGACCTGTCTTTATTTAGCCAGGTGTTTTAT 600
QY 2727 TACCAAGATGTTTAAAGTGCAGAGAGAAATGTATGATAAGATATCATGCTTCAGATT 2786
Db 601 TACCAAGATGTTTAAAGTGCAGAGAGAAATGTATGATAAGATATCATGCTTCAGATT 660
QY 2787 GGTGCATCTTTAATGGATCCCAATTAAGTTCTTGTGTCTGTACTTTCAGAGGTATGAATT 2846
Db 661 GGTGCATCTTTAATGGATCCCAATTAAGTTCTTGTGTCTGTACTTTCAGAGGTATGAATT 720
QY 2847 GCCGAGGCTTTTAAACAGACCATATCTACAAAGACACAGGATTTGATTTAAACAATATAAT 2906
Db 721 GCCGAGGCTTTTAAACAGACCATATCTACAAAGACACAGGATTTGATTTAAACAATATAAT 780
QY 2907 ACATTAATAGAGAAATGCTTCAGGTCTCTCATCTATATTGTGGGTGAGGTTTATGTACT 2966
Db 781 ACATTAATAGAGAAATGCTTCAGGTCTCTCATCTATATTGTGGGTGAGGTTTATGTACT 840
QY 2967 GGAGTGGGAAATGTGACCAAGAGAGGTCAATGAGAGAAATCATTCATTGCTTTTGC 3026
Db 841 GGAGTGGGAAATGTGACCAAGAGAGGTCAATGAGAGAAATCATTCATTGCTTTTGC 900
QY 3027 ATTGAACCCATGCCACAGTGCCTATGCCAAATTTTACCTGAGAAATGAAATATAATGAA 3086
Db 901 ATTGAACCCATGCCACAGTGCCTATGCCAAATTTTACCTGAGAAATGAAATATAATGAA 960
QY 3087 ACTGCTTTAGAGAAATGTATAAACAAAGTGGCCACATTTAAGAAACACAGGTGTATCAGGC 3146
Db 961 ACTGCTTTAGAGAAATGTATAAACAAAGTGGCCACATTTAAGAAACACAGGTGTATCAGGC 1020
QY 3147 CATGAGTTTATGAATCTAAAGATGAATCACTGAAAGACTTCAATATGTACTTTTATCAT 3206
Db 1021 CATGAGTTTATGAATCTAAAGATGAATCACTGAAAGACTTCAATATGTACTTTTATCAT 1080
QY 3207 TACTTCAAAACCCAGCATAGCAGGCTGAAATATGAGAGAAAGAGAGAAACAAAGAA 3266
Db 1081 TACTTCAAAACCCAGCATAGCAGGCTGAAATATGAGAGAAAGAGAGAAACAAAGAA 1140
QY 3267 AACAAAGATGAAGCATTTGCCGCCACACACCTCTGTAATTTCTGCCCTGCTTTTACAGAAA 3326
Db 1141 AACAAAGATGAAGCATTTGCCGCCACACACCTCTGTAATTTCTGCCCTGCTTTTACAGAAA 1200
QY 3327 GTGATTAACTTCTCAACTGTGTATCATGATGTATCTTCTCAGGACCGTATTTGAGCGG 3386

Db 1201 GTGATTAACTTCTCAACTGTGTATCATGATGTATCTTCTCAGGACCGTATTTGAGCGG 1260
QY 3387 GCAATAGACACAGATTCTTAACCTTGTGGACCGAAGGGATGCTCCAATGGCTTTTCATATT 3446
Db 1261 GCAATAGACACAGATTCTCAACTTGTGGACCGAAGGGATGCTCCAATGGATTGGGAGTT 1320
QY 3447 CTGGCAT 3453
Db 1321 CAGCCAT 1327

Search completed: July 8, 2005, 19:42:33
Job time : 2927 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 8, 2005, 18:46:00 ; Search time 993 Seconds
(without alignments)
10394.394 Million cell updates/sec

Title: US-10-758-672A-1

Perfect score: 6308

Sequence: 1 gccaagattcgccagcagg.....aattttgtattgtgtttt 6308

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PTUS COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6308	100.0	6308	4	US-09-724-126A-1
2	5136	81.4	5205	4	US-09-724-126A-18
3	4435.4	70.3	6395	2	US-08-982-956-1
4	4435.4	70.3	6395	3	US-09-228-317-1
5	1381.8	21.9	6300	4	US-09-724-126A-3
6	1233.6	19.6	6089	4	US-09-724-126A-5
7	996.2	15.8	1001	2	US-08-982-956-2
8	996.2	15.8	1001	3	US-09-228-317-2
9	113.4	1.7	670	4	US-09-270-767-14761
10	105.6	1.7	366	4	US-09-513-999C-936
11	105.6	1.7	456	4	US-09-513-999C-938
12	103.2	1.6	605	4	US-09-513-999C-937
13	100.8	1.6	602	4	US-09-513-999C-14808
14	100.8	1.6	692	4	US-09-513-999C-14809
15	100.8	1.6	842	4	US-09-513-999C-14947
16	85.2	1.4	2182	1	US-09-270-767-13447
17	57.4	0.9	7218	1	US-08-232-463-14
18	52.2	0.8	7218	1	US-08-232-463-14
19	50.4	0.8	601	4	US-09-949-016-179951
20	50.4	0.8	103356	4	US-09-949-016-12364
21	50.4	0.8	101357	4	US-09-949-016-16924
22	49.2	0.8	932	4	US-09-573-080A-76
23	49	0.8	1141	4	US-09-806-708B-22
24	47.4	0.8	601	4	US-09-949-016-86450
25	47	0.7	128999	4	US-09-949-016-14684
26	47	0.7	148156	4	US-09-949-016-11776
27	46.2	0.7	95648	4	US-09-949-016-13139

28	46.2	0.7	183112	4	US-09-949-016-14184	Sequence 14184, A
29	46	0.7	43333	4	US-09-949-016-15381	Sequence 15381, A
30	46	0.7	300402	4	US-09-949-016-13632	Sequence 13632, A
31	45.8	0.7	601	4	US-09-949-016-70752	Sequence 70752, A
32	45.8	0.7	601	4	US-09-949-016-70753	Sequence 70753, A
33	45.8	0.7	601	4	US-09-949-016-70754	Sequence 70754, A
34	45.8	0.7	106924	4	US-09-949-016-13834	Sequence 13834, A
35	45.6	0.7	601	4	US-09-949-016-69710	Sequence 69710, A
36	45.6	0.7	601	4	US-09-949-016-147873	Sequence 147873, A
37	45.6	0.7	72455	4	US-09-949-016-13793	Sequence 13793, A
38	45.6	0.7	76472	4	US-09-949-016-15896	Sequence 15896, A
39	45.6	0.7	92276	4	US-09-949-016-12166	Sequence 12166, A
40	45.6	0.7	104475	4	US-09-949-016-12115	Sequence 12115, A
41	45.6	0.7	111282	3	US-09-754-250-3	Sequence 3, Appli
42	45.2	0.7	601	4	US-09-949-016-168631	Sequence 168631, A
43	45.2	0.7	601	4	US-09-949-016-174776	Sequence 174776, A
44	45.2	0.7	5932	4	US-09-949-016-14671	Sequence 14671, A
45	45.2	0.7	15720	4	US-09-949-016-11780	Sequence 11780, A

ALIGNMENTS

RESULT 1

US-09-724-126A-1

; Sequence 1, Application US/09724126A

; Patent No. 6706505

; GENERAL INFORMATION:

; APPLICANT: Han, Hui-Quan

; APPLICANT: Kwak, Keith

; TITLE OF INVENTION: Human E3 Alpha Ubiquitin Ligase Family

; FILE REFERENCE: 01017/35966A

; CURRENT APPLICATION NUMBER: US/09/724.126A

; CURRENT FILING DATE: 2000-11-28

; PRIOR APPLICATION NUMBER: US 60/187,211

; PRIOR FILING DATE: 1999-03-01

; NUMBER OF SEQ ID NOS: 29

; SOFTWARE: Patent in version 3.0

; SEQ ID NO 1

; LENGTH: 6308

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (696)..(5942)

; US-09-724-126A-1

Query Match 100.0%; Score 6308; DB 4; Length 6308;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 6308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GCCAAGAAATTCGGCACGAGGGGAAAAGCTGAGCCGAGGAACCAAAATTAATTTGCTTTACT	60
Db	1	GCCAAGAAATTCGGCACGAGGGGAAAAGCTGAGCCGAGGAACCAAAATTAATTTGCTTTACT	60
Qy	61	CATTGTGTAGACAAGCGTCAAAAACAGCTTCACCTATCTTGAACAAGAACTTACT	120
Db	61	CATTGTGTAGACAAGCGTCAAAAACAGCTTCACCTATCTTGAACAAGAACTTACT	120
Qy	121	CCAAAGGCTTATCATCTGTCTTCCACTTATCCACNAGCTGCTATGGCAGCTGCTGTC	180
Db	121	CCAAAGGCTTATCATCTGTCTTCCACTTATCCACNAGCTGCTATGGCAGCTGCTGTC	180
Qy	181	CGCACCTGGAAACACCGCCAGCCCTACTGCTCCACTACCTACCTGCTTCTCCACCT	240
Db	181	CGCACCTGGAAACACCGCCAGCCCTACTGCTCCACTACCTGCTTCTCCACCT	240
Qy	241	GATCAGCTGCTTGTGCTGCCATCTTATCCGCTTCTGCTGCTTCTGAGTAAATGTATACA	300
Db	241	GATCAGCTGCTTGTGCTGCCATCTTATCCGCTTCTGCTGCTTCTGAGTAAATGTATACA	300
Qy	301	CACCTGGAAACACCACTTCTTCTGCTGCTTCTGCTGCTTCTGCTGCTTCTGCTGCTG	360
Db	301	CACCTGGAAACACCACTTCTTCTGCTGCTTCTGCTGCTTCTGCTGCTTCTGCTGCTG	360

Db 301 CACCCTGGAAACCACTTACTTCTGTGTCTATGAATTTGACTACTCTAGCTGGATC 360
Qy 361 CCGAGCTTTTGTATACATGTGCAAGTGGCCACGGGTAGAACTCTTAATAATAGAAT 420
Db 361 CCGAGCTTTTGTATACATGTGCAAGTGGCCACGGGTAGAACTCTTAATAATAGAAT 420
Qy 421 GTATGCAACAGTTTCCAGCACCAACCCAGATATACAAACCAATTCAGCTTACCAAGAGCTAC 480
Db 421 GTATGCAACAGTTTCCAGCACCAACCCAGATATACAAACCAATTCAGCTTACCAAGAGCTAC 480
Qy 481 GCCTGATAAATTTAGAGGGGAAAAAATCTCCAGTCCCTTCAGTCTGTCAGCTGCT 540
Db 481 GCCTGATAAATTTAGAGGGGAAAAAATCTCCAGTCCCTTCAGTCTGTCAGCTGCT 540
Qy 541 TCCGGGAAGCGGCGCGGAAGCAGTCTCTCCGACCGGCCACAGGTTTCGGCTTCCCTGCGGCG 600
Db 541 TCCGGGAAGCGGCGCGGAAGCAGTCTCTCCGAGTCTGCGTCAAAACCCGACTTCAGGGGCG 600
Qy 601 TCGTAAAAAGTGTGCTCTCTCCGACCGGCCACAGGTTTCGGCTTCCCTGCGGCG 660
Db 601 TCGTAAAAAGTGTGCTCTCTCCGACCGGCCACAGGTTTCGGCTTCCCTGCGGCG 660
Qy 661 GGGGTCCGCAACTGCAAGCGTCAAGTTCCTCAAGATGCGGACGAGGAGGCTGGAGTA 720
Db 661 GGGGTCCGCAACTGCAAGCGTCAAGTTCCTCAAGATGCGGACGAGGAGGCTGGAGTA 720
Qy 721 CTGAGAGATGGAATCAGCGGAGTACCCGACGCCCTCAGGCTCGGATCTGGT 780
Db 721 CTGAGAGATGGAATCAGCGGAGTACCCGACGCCCTCAGGCTCGGATCTGGT 780
Qy 781 GGGATCAGCAAGTTGATTTTATCTCTCTCCGATCATTTGGGCAAAATTTGGTCCAG 840
Db 781 GGGATCAGCAAGTTGATTTTATCTCTCTCTCCGATCATTTGGGCAAAATTTGGTCCAG 840
Qy 841 AAATTTACTTGTCTGAAATGGACCCAGACTTTGGAAAGCAGGAGAAAGTGTACAAATGT 900
Db 841 AAATTTACTTGTCTGAAATGGACCCAGACTTTGGAAAGCAGGAGAAAGTGTACAAATGT 900
Qy 901 CAATATTCATCTGGAATGGTACTTATTTGGAGAGATCCAGATATTTGCTTAGAGA 960
Db 901 CAATATTCATCTGGAATGGTACTTATTTGGAGAGATCCAGATATTTGCTTAGAGA 960
Qy 961 AATTGAACACAGTGGAGCATTTACGCTTTGGGAGGGTTTTCAAAAGTGGAGAGACA 1020
Db 961 AATTGAACACAGTGGAGCATTTACGCTTTGGGAGGGTTTTCAAAAGTGGAGAGACA 1020
Qy 1021 CCTATTCTTGCAAGGATTTGCAATTTGATCCAAATGTTACTCTGTATGGAAGTCTTCC 1080
Db 1021 CCTATTCTTGCAAGGATTTGCAATTTGATCCAAATGTTACTCTGTATGGAAGTCTTCC 1080
Qy 1081 AGGACAGTGTTCATAAAAATCATGTTACAGATGCATCTTACTTGGAGGGGTTCT 1140
Db 1081 AGGACAGTGTTCATAAAAATCATGTTACAGATGCATCTTACTTGGAGGGGTTCT 1140
Qy 1141 GTGACTGTGGAGACAGAGGATCGAAAACTGGCCCTTTTGTGTTAAATCATGAACCTG 1200
Db 1141 GTGACTGTGGAGACAGAGGATCGAAAACTGGCCCTTTTGTGTTAAATCATGAACCTG 1200
Qy 1201 GAAAGACAGGTACTATAAAGAGAAATTCACGCTGTCGGTTGAAATGAAGAGTAAATGTC 1260
Db 1201 GAAAGACAGGTACTATAAAGAGAAATTCACGCTGTCGGTTGAAATGAAGAGTAAATGTC 1260
Qy 1261 AAGCAGGAAAAATATTTCTTCAAGTGAATAAATATGTCGTGAATAATGATATGCGGAAG 1320
Db 1261 AAGCAGGAAAAATATTTCTTCAAGTGAATAAATATGTCGTGAATAATGATATGCGGAAG 1320
Qy 1321 AGGAAAAAGAACTGCTCTGAACTCCAGATAAGGAGAGAAAAATCAAGAGATACTTGTG 1380
Db 1321 AGGAAAAAGAACTGCTCTGAACTCCAGATAAGGAGAGAAAAATCAAGAGATACTTGTG 1380
Qy 1381 TCCTTTTCAATGATGAACCACTTATGATGACCGTCAATATACAGCCCTACAAAGAGCTC 1440
Db 1381 TCCTTTTCAATGATGAACCACTTATGATGACCGTCAATATACAGCCCTACAAAGAGCTC 1440

Qy 1441 TTGACTGTGAGCTCGCAGAGSCCAGTTTGATACCACTGCCATTTGACAAAGAGGGTGGTC 1500
Db 1441 TTGACTGTGAGCTCGCAGAGSCCAGTTTGATACCACTGCCATTTGACAAAGAGGGTGGTC 1500
Qy 1501 GGGCTGTTAAAGCGGGAGCTTATGCTGCTTCCAGGAAGCAAGAGATATAAAGAGTC 1560
Db 1501 GGGCTGTTAAAGCGGGAGCTTATGCTGCTTCCAGGAAGCAAGAGATATAAAGAGTC 1560
Qy 1561 ATTGAGAAAAATGCTCTCAACATCCATTCATGTAGAGTATTACATCTCAGAGATTATGG 1620
Db 1561 ATTGAGAAAAATGCTCTCAACATCCATTCATGTAGAGTATTACATCTCAGAGATTATGG 1620
Qy 1621 CTCAATCAGAAAAATTTGCTTGTGCTCTGCTCTGATGAAACAAATTTATGAGCTATTCAA 1680
Db 1621 CTCAATCAGAAAAATTTGCTTGTGCTCTGCTCTGATGAAACAAATTTATGAGCTATTCAA 1680
Qy 1681 GTGACTTTTAGGCAGATCTTTTCCAGCATGCCCTTAGAGAGAAACCTGACTCGGAGATC 1740
Db 1681 GTGACTTTTAGGCAGATCTTTTCCAGCATGCCCTTAGAGAGAAACCTGACTCGGAGATC 1740
Qy 1741 CCTGTCTCATAGCAGGTTAATGCTTTGGGATGCAAGCTTTTAAAGGTGCCCGTAAGA 1800
Db 1741 CCTGTCTCATAGCAGGTTAATGCTTTGGGATGCAAGCTTTTAAAGGTGCCCGTAAGA 1800
Qy 1801 TCCTTTCATGAATTCATCTTCAGCAGATTTTATATGGAGATGGAATACAAAAAACTCTTTG 1860
Db 1801 TCCTTTCATGAATTCATCTTCAGCAGATTTTATATGGAGATGGAATACAAAAAACTCTTTG 1860
Qy 1861 CTATGGAATTTGTGAAGTATTATAACAACCTGCAAGAAATATATCATGATGATCATG 1920
Db 1861 CTATGGAATTTGTGAAGTATTATAACAACCTGCAAGAAATATATCATGATGATCATG 1920
Qy 1921 ACAGAGTATCTCTATAACTGCATCTTCAGTTCAGATGTTTACTGTTCTTACTCTGGCTC 1980
Db 1921 ACAGAGTATCTCTATAACTGCATCTTCAGTTCAGATGTTTACTGTTCTTACTCTGGCTC 1980
Qy 1981 GACATCTTATTTGAAGAGCAGAACTGTTATCTCTGCTTACTTACTGTTTCTTCTCTCT 2040
Db 1981 GACATCTTATTTGAAGAGCAGAACTGTTATCTCTGCTTACTTACTGTTTCTTCTCTCT 2040
Qy 2041 TACCTGAGTACTTGGACAGGAACAAATAAATTCACCTCCAGGGTTATAGCCAGGACAAAT 2100
Db 2041 TACCTGAGTACTTGGACAGGAACAAATAAATTCACCTCCAGGGTTATAGCCAGGACAAAT 2100
Qy 2101 TGGGAAGATATATGCAAGTATATGTAACCTTAAAGTATATCTGATCAGCAAAACCCACA 2160
Db 2101 TGGGAAGATATATGCAAGTATATGTAACCTTAAAGTATATCTGATCAGCAAAACCCACA 2160
Qy 2161 TATGACACAGAAAGATTAAAGATGCAATTCCTTGAAGGTTTTCGATCTTTTGAAGATTC 2220
Db 2161 TATGACACAGAAAGATTAAAGATGCAATTCCTTGAAGGTTTTCGATCTTTTGAAGATTC 2220
Qy 2221 TTACTCTGTATGCAAGGAAATGGAAGAAATCCGAAGACAGGTTGGGCAACATTTGAAGTGG 2280
Db 2221 TTACTCTGTATGCAAGGAAATGGAAGAAATCCGAAGACAGGTTGGGCAACATTTGAAGTGG 2280
Qy 2281 ATCTGTATGGAAGGCTGCCATTCGTATACAGATGCAATTCGAAGATTTTACTCATGT 2340
Db 2281 ATCTGTATGGAAGGCTGCCATTCGTATACAGATGCAATTCGAAGATTTTACTCATGT 2340
Qy 2341 TCCAAGAGTGTGTGCTTGTGATGAAGAACTCTTACTTGTGCTTATATAAGAAATGTCCACA 2400
Db 2341 TCCAAGAGTGTGTGCTTGTGATGAAGAACTCTTACTTGTGCTTATATAAGAAATGTCCACA 2400
Qy 2401 AAGCTGTGATGAGGTGCAAGTACCAAGTTCATATCTAGTAGCAAGACAGTAGTCAATCGT 2460
Db 2401 AAGCTGTGATGAGGTGCAAGTACCAAGTTCATATCTAGTAGCAAGACAGTAGTCAATCGT 2460
Qy 2461 GTGACATAGTTTGGAAAAAAGTCTTACAGATCTCTGAGGATCTTGTAGGATACATC 2520
Db 2461 GTGACATAGTTTGGAAAAAAGTCTTACAGATCTCTGAGGATCTTGTAGGATACATC 2520

QY 2521 TGCACCTCTAGGACCCCTTGTGCTCTTCATGTAGCTTTAAGCAGGCTGGGTGCTGTTT 2580
Db 2521 TGCACCTCTAGGACCCCTTGTGCTCTTCATGTAGCTTTAAGCAGGCTGGGTGCTGTTT 2580
QY 2581 CAAGACTGCATGAATTTGTGCTCTTTAGGACTTTCAAGTAGAGGTACTAGTGGAAATATC 2640
Db 2581 CAAGACTGCATGAATTTGTGCTCTTTAGGACTTTCAAGTAGAGGTACTAGTGGAAATATC 2640
QY 2641 CTTTACGTTGCTGTGTGTGTTGCCAGGTTTGTCTGAGATGTGGCGAAGAAATGGAC 2700
Db 2641 CTTTACGTTGCTGTGTGTGTTGCCAGGTTTGTCTGAGATGTGGCGAAGAAATGGAC 2700
QY 2701 TGTCTCTTATAGCCAGGTGTTTATACCAAGATGTTAAGTGCAGAGAGAAATGTATG 2760
Db 2701 TGTCTCTTATAGCCAGGTGTTTATACCAAGATGTTAAGTGCAGAGAGAAATGTATG 2760
QY 2761 ATAAAGATATCATCATGCTTCAGATGTGTGCACTTTTAATGATCCCAATAGTCTTGT 2820
Db 2761 ATAAAGATATCATCATGCTTCAGATGTGTGCACTTTTAATGATCCCAATAGTCTTGT 2820
QY 2821 TACTGGTACTTCAGAGGTATGAATTTGCCAGGCTTTTAAACAAGACCATATCTACAAAAG 2880
Db 2821 TACTGGTACTTCAGAGGTATGAATTTGCCAGGCTTTTAAACAAGACCATATCTACAAAAG 2880
QY 2881 ACCAGGATTTGATTAACAATATAATATACATTAATAGAGAAATGTTCAGGTCTCATCT 2940
Db 2881 ACCAGGATTTGATTAACAATATAATATACATTAATAGAGAAATGTTCAGGTCTCATCT 2940
QY 2941 ATATTGTGGGTGAGGTTATGTACCTGGAGTGGGAAATGTACCAAGAGAGGTACAA 3000
Db 2941 ATATTGTGGGTGAGGTTATGTACCTGGAGTGGGAAATGTACCAAGAGAGGTACAA 3000
QY 3001 TGAGAGAAATCAATTCACCTTGTCTTGTGCAATGCAACCAATGCCACACAGTGCCTCAATA 3060
Db 3001 TGAGAGAAATCAATTCACCTTGTCTTGTGCAATGCAACCAATGCCACACAGTGCCTCAATA 3060
QY 3061 ATTTACCTGAGAAATGAATTAATGAATCTGGCTTAGAGAAATGTATAAACAAAGTGGCCA 3120
Db 3061 ATTTACCTGAGAAATGAATTAATGAATCTGGCTTAGAGAAATGTATAAACAAAGTGGCCA 3120
QY 3121 CATTTAAGAAACCAAGTGTATCAGGCCATGGAGTTTATGAATGAATGAATCACTGA 3180
Db 3121 CATTTAAGAAACCAAGTGTATCAGGCCATGGAGTTTATGAATGAATGAATCACTGA 3180
QY 3181 AAGACTTCAATATGTACTTTTATCAATTAATCTCAAAACCCAGCATAGCAGGCTGAACATA 3240
Db 3181 AAGACTTCAATATGTACTTTTATCAATTAATCTCAAAACCCAGCATAGCAGGCTGAACATA 3240
QY 3241 TGCAGAGAAAGGAGAAACCAAGAAACCAAGATGAAGCATTTGCCGCCACCAACCTC 3300
Db 3241 TGCAGAGAAAGGAGAAACCAAGAAACCAAGATGAAGCATTTGCCGCCACCAACCTC 3300
QY 3301 CTGAATTCGCTGCTTTTCAGCAAGTGAATTAACCTTCTCACTGTGATATCATGATGT 3360
Db 3301 CTGAATTCGCTGCTTTTCAGCAAGTGAATTAACCTTCTCACTGTGATATCATGATGT 3360
QY 3361 ACATTCTCAGGACCGTATTTGAGCGGCAATAGACACAGATTTCACTTTGTGGACCGAAG 3420
Db 3361 ACATTCTCAGGACCGTATTTGAGCGGCAATAGACACAGATTTCACTTTGTGGACCGAAG 3420
QY 3421 GGATGCTCCAAATGGCTTTTATATTTCTGGCATTTGGTGTATAGAGAGAGCAACAGC 3480
Db 3421 GGATGCTCCAAATGGCTTTTATATTTCTGGCATTTGGTGTATAGAGAGAGCAACAGC 3480
QY 3481 TTCAAAAGCTTCTGAAGAGAGTAAATTTGACTTTTATCATAGGCTTCAAGATTTGG 3540
Db 3481 TTCAAAAGCTTCTGAAGAGAGTAAATTTGACTTTTATCATAGGCTTCAAGATTTGG 3540
QY 3541 GAAGTTTCAGGCATGAATATACAAATGCTTTTGGAAATCTCAAGGAAATTTCCCAAGTTAG 3600
Db 3541 GAAGTTTCAGGCATGAATATACAAATGCTTTTGGAAATCTCAAGGAAATTTCCCAAGTTAG 3600
QY 3601 AAGGCCAGAGGACATGATAACGTGGATACCTTCAGATGTTTGAACACAGTGAAGCGATTAA 3660

Db 3601 AAGGCCAGAGGACATGATAACGTGGATACCTTCAGATGTTTGAACACAGTGAAGCGATTAA 3660
QY 3661 GAGAAAAATCTTGTATTTAGTGTAGCAACCATCAGGATCGGAATCTATTAAGAATGATG 3720
Db 3661 GAGAAAAATCTTGTATTTAGTGTAGCAACCATCAGGATCGGAATCTATTAAGAATGATG 3720
QY 3721 AGATTACTCATGATAAAGAAAAAGCAGAAACGAAAAAGAGCTGAAAGCTGTAGGCTAC 3780
Db 3721 AGATTACTCATGATAAAGAAAAAGCAGAAACGAAAAAGAGCTGAAAGCTGTAGGCTAC 3780
QY 3781 ATCCCGAAGATCATGCTCAGATGTCTGCTTACAGAAAACTTCATTGAAACTCATA 3840
Db 3781 ATCCCGAAGATCATGCTCAGATGTCTGCTTACAGAAAACTTCATTGAAACTCATA 3840
QY 3841 AACTCATGATATGACAAATACATCAGAAATGCTGGGAAAGAGATTTCCATTTAGAGGAAG 3900
Db 3841 AACTCATGATATGACAAATACATCAGAAATGCTGGGAAAGAGATTTCCATTTAGAGGAAG 3900
QY 3901 AGAGACCCCAAGCAGTCACTCTAGAAATGCTTTGGGTCTTAAACGGGGTCCAT 3960
Db 3901 AGAGACCCCAAGCAGTCACTCTAGAAATGCTTTGGGTCTTAAACGGGGTCCAT 3960
QY 3961 CTGTTACTGAAAAGAGGTGCTGATGCTCTTGGCCAGAGAAACAGGAGGTGAAA 4020
Db 3961 CTGTTACTGAAAAGAGGTGCTGATGCTCTTGGCCAGAGAAACAGGAGGTGAAA 4020
QY 4021 TAGAAAAATATGCAATGTTATTTATCGGCTCTGCTCCAGAAATCTACTGCCCTTAAACCCAGC 4080
Db 4021 TAGAAAAATATGCAATGTTATTTATCGGCTCTGCTCCAGAAATCTACTGCCCTTAAACCCAGC 4080
QY 4081 ACAGGGGAAAAACCATAGAACTCTCAGAGAAAGCCCTAGACCCACTTTTTCATGATCCAG 4140
Db 4081 ACAGGGGAAAAACCATAGAACTCTCAGAGAAAGCCCTAGACCCACTTTTTCATGATCCAG 4140
QY 4141 ACTTGGCATATGGAATTTATACAGGAGTGTGCTCATGTAATGCAACGAGTGTCTGGC 4200
Db 4141 ACTTGGCATATGGAATTTATACAGGAGTGTGCTCATGTAATGCAACGAGTGTCTGGC 4200
QY 4201 AGAAGTATTTTGAAGCTGTACAGCTCTCAGCAGCGCATTCATGTCACCTTTTGTG 4260
Db 4201 AGAAGTATTTTGAAGCTGTACAGCTCTCAGCAGCGCATTCATGTCACCTTTTGTG 4260
QY 4261 ACTTGGAAAGTGGAGAAATCTTTGCCCTCTTTGCAATCTCTGTGCAATCTGTGATCC 4320
Db 4261 ACTTGGAAAGTGGAGAAATCTTTGCCCTCTTTTGGCAATCTCTGTGCAATCTGTGATCC 4320
QY 4321 CCATTATTTCTTTCGAACCTCAAAAGATAAACAGTGAAGTGCAGATGCTCTTGTCTAAC 4380
Db 4321 CCATTATTTCTTTCGAACCTCAAAAGATAAACAGTGAAGTGCAGATGCTCTTGTCTAAC 4380
QY 4381 TTTTGAACCTGGCAGCTGTGATACAGCTGTTCTGGCCAGAAATATCAGGTTTATATATA 4440
Db 4381 TTTTGAACCTGGCAGCTGTGATACAGCTGTTCTGGCCAGAAATATCAGGTTTATATATA 4440
QY 4441 GACATGCTTAAAGGAGAAAAACCAATTCCTATTTCTTTAAATCAAGGATGGAGATTTCTA 4500
Db 4441 GACATGCTTAAAGGAGAAAAACCAATTCCTATTTCTTTAAATCAAGGATGGAGATTTCTA 4500
QY 4501 CTTTGGAGTTCCATTCATCTGAGTTTGGGCTGTGATCTTCGATTAATATTTCAATA 4560
Db 4501 CTTTGGAGTTCCATTCATCTGAGTTTGGGCTGTGATCTTCGATTAATATTTCAATA 4560
QY 4561 GCATCAAGGAAATGTTATTTCTTTTGCACAAATTTATAGAAATGGATTGAAAGTGC 4620
Db 4561 GCATCAAGGAAATGTTATTTCTTTTGCACAAATTTATAGAAATGGATTGAAAGTGC 4620
QY 4621 CACCTGATGAAGAGGATCTCGAGTCCCTCAGCTGAGCCTGAGGACCTGCGCTTTCACTA 4680
Db 4621 CACCTGATGAAGAGGATCTCGAGTCCCTCAGCTGAGCCTGAGGACCTGCGCTTTCACTA 4680
QY 4681 TCCAGGCAATTTGAAAAATCTATTGGGAGATGAAGAAAAACCTCTGTGTTGGAGCATTTCAA 4740

Db 4681 TCCAGGCAATGAAAATCTATTGGAGATGAAGGAAAAACCTCTGTTTGGAGCACTTCAAA 4740
Qy 4741 ATAGCAGCATATGCTGTGAAGCAATTAATGCAAGTTTGCAGTTTGCACAGAGGATTACCT 4800
Db 4741 ATAGCAGCATATGCTGTGAAGCAATTAATGCAAGTTTGCAGTTTGCACAGAGGATTACCT 4800
Qy 4801 GTCCCTCAGTCTCTGATACAGAAACATCTGGTTTGGTCTTCTATCAGTCTGTTCTTCTTAACA 4860
Db 4801 GTCCCTCAGTCTCTGATACAGAAACATCTGGTTTGGTCTTCTATCAGTCTGTTCTTCTTAACA 4860
Qy 4861 TAAATACAGAGATACACATGCTCTCTCTATAGATCTGTTTCAATGTTTGGTGGTG 4920
Db 4861 TAAATACAGAGATACACATGCTCTCTCTATAGATCTGTTTCAATGTTTGGTGGTG 4920
Qy 4921 CTGTTTGGATCTCCATCTCTGTTTGGATGACCTGTTGATCTGAGCCCTTCTTCAG 4980
Db 4921 CTGTTTGGATCTCCATCTCTGTTTGGATGACCTGTTGATCTGAGCCCTTCTTCAG 4980
Qy 4981 TTAGTTCTTCTTATACACACCTTTATCTTCTTCCATTTGATCACCATGGCACATGCTTC 5040
Db 4981 TTAGTTCTTCTTATACACACCTTTATCTTCTTCCATTTGATCACCATGGCACATGCTTC 5040
Qy 5041 AGATACTACTTACAGTAGACACAGCCCTACCCCTTGTCTCAGTTTCAAGAAGACAGTGAAG 5100
Db 5041 AGATACTACTTACAGTAGACACAGCCCTACCCCTTGTCTCAGTTTCAAGAAGACAGTGAAG 5100
Qy 5101 AGGCTCATCTCGCATCTCTTCTTCTTGGAGAAATTTCTCAATATACAAGTGGCTCCATTG 5160
Db 5101 AGGCTCATCTCGCATCTCTTCTTCTTGGAGAAATTTCTCAATATACAAGTGGCTCCATTG 5160
Qy 5161 GGTGGATATCTCGCTGGTATTTGTTGGTCTCACTGAAGAAATGGCATACCCCTTATC 5220
Db 5161 GGTGGATATCTCGCTGGTATTTGTTGGTCTCACTGAAGAAATGGCATACCCCTTATC 5220
Qy 5221 TTGCTGTGCTGCAATGTTTTTCCATTTTCACTATTATCTTGGGGTAACTCCGCTCAGGAACTGC 5280
Db 5221 TTGCTGTGCTGCAATGTTTTTCCATTTTCACTATTATCTTGGGGTAACTCCGCTCAGGAACTGC 5280
Qy 5281 ATACCAATTTGACAGAGAGTACAGTGCATCTCTGTAGCTATCTATCTTACCTACAA 5340
Db 5281 ATACCAATTTGACAGAGAGTACAGTGCATCTCTGTAGCTATCTATCTTACCTACAA 5340
Qy 5341 ATTTGTTCTGCTCTTCCAGGAATATTTGGATACGTAAAGGCCCTTGTCTCCAGAGTGGT 5400
Db 5341 ATTTGTTCTGCTCTTCCAGGAATATTTGGATACGTAAAGGCCCTTGTCTCCAGAGTGGT 5400
Qy 5401 GTGCAGATCTGCTCTTACTTAACTGTTTGAAGCAAAAAACACCGTGTGAGTACCTTA 5460
Db 5401 GTGCAGATCTGCTCTTACTTAACTGTTTGAAGCAAAAAACACCGTGTGAGTACCTTA 5460
Qy 5461 GAAAAAGAAATAGTTTGTATAGAGCTTCTGTATAGCTATAGCTGCTCTGTAATCAAGCTT 5520
Db 5461 GAAAAAGAAATAGTTTGTATAGAGCTTCTGTATAGCTATAGCTGCTCTGTAATCAAGCTT 5520
Qy 5521 CTCATTTCAGGTCGCCAGCTCTGAGATGATGAGCAAGCATCTGTCTCTGCTCTT 5580
Db 5521 CTCATTTCAGGTCGCCAGCTCTGAGATGATGAGCAAGCATCTGTCTCTGCTCTT 5580
Qy 5581 TCTGTGGGCTACTACTATGTTCTCAGAACATTTGCTGCCAGGAAATTTGTAACGGGGAAG 5640
Db 5581 TCTGTGGGCTACTACTATGTTCTCAGAACATTTGCTGCCAGGAAATTTGTAACGGGGAAG 5640
Qy 5641 AGTTTGGAGCTTGCAATTTTTCAGGCACTTCACTGTGGAGCCGGAGTCTGCAATTTCTTAA 5700
Db 5641 AGTTTGGAGCTTGCAATTTTTCAGGCACTTCACTGTGGAGCCGGAGTCTGCAATTTCTTAA 5700
Qy 5701 AAATCAGAGAAATGCGAGTGTCTGTTGAGGTAAGCCAGAGGCTGTGCTATCCAG 5760
Db 5701 AAATCAGAGAAATGCGAGTGTCTGTTGAGGTAAGCCAGAGGCTGTGCTATCCAG 5760
Qy 5761 CTCCTTACTTGGATGAATATGAGAAACAGACCTCTGGCTGAAGAGGGGCAACCCCTTTC 5820
Db 5761 CTCCTTACTTGGATGAATATGAGAAACAGACCTCTGGCTGAAGAGGGGCAACCCCTTTC 5820

Qy 5821 ATTTATCTCGTACGGGTATCGAAGCTCCATTTGGTCTGCGCAACACACCTGCATTATAG 5880
Db 5821 ATTTATCTCGTACGGGTATCGAAGCTCCATTTGGTCTGCGCAACACACCTGCATTATAG 5880
Qy 5881 AAGAGATTGCTAGAGCAAGAGACTAATCAGATGTTTATTTGGATTCAACTGGCAGTTTAC 5940
Db 5881 AAGAGATTGCTAGAGCAAGAGACTAATCAGATGTTTATTTGGATTCAACTGGCAGTTTAC 5940
Qy 5941 TGTGAGCTCCAACTCTGCTCAAGCAATCAAAATGACGACAGTAGTAAAGGCTGATTC 6000
Db 5941 TGTGAGCTCCAACTCTGCTCAAGCAATCAAAATGACGACAGTAGTAAAGGCTGATTC 6000
Qy 6001 AAAATTTATGGAAGAACTTTCTGAGGGCTGGGAAAGTATTGGAGGGTCTTTTCTCCATGTC 6060
Db 6001 AAAATTTATGGAAGAACTTTCTGAGGGCTGGGAAAGTATTGGAGGGTCTTTTCTCCATGTC 6060
Qy 6061 CAGGTTCACTTACATCAATAAATTTTCTTAAATGGAGTATTGCTTTCAATTAGCAACA 6120
Db 6061 CAGGTTCACTTACATCAATAAATTTTCTTAAATGGAGTATTGCTTTCAATTAGCAACA 6120
Qy 6121 TATGCTTCAAGGAAAAAGGACATAGATCAATCTGTTTATGCTAGTATTTCCAGGA 6180
Db 6121 TATGCTTCAAGGAAAAAGGACATAGATCAATCTGTTTATGCTAGTATTTCCAGGA 6180
Qy 6181 ATTTATTTCCCTTCCATAAATTTTCTCAATTTTATTTTCACTTGGTAGATGAAG 6240
Db 6181 ATTTATTTCCCTTCCATAAATTTTCTCAATTTTATTTTCACTTGGTAGATGAAG 6240
Qy 6241 TCAGGTCAAACAGTTGTAGACATTTATGTTTGGTTAACTCTTCTGCAATTTTGTATTT 6300
Db 6241 TCAGGTCAAACAGTTGTAGACATTTATGTTTGGTTAACTCTTCTGCAATTTTGTATTT 6300
Qy 6301 GGTGTTTT 6308
Db 6301 GGTGTTTT 6308

RESULT 2

US-09-724-126A-18
; Sequence 18, Application US/09724126A
; Patent No. 6706505
; GENERAL INFORMATION:
; APPLICANT: Han, Hui-Quan
; APPLICANT: Kwak, Keich
; TITLE OF INVENTION: Human E3 Alpha Ubiquitin Ligase Family
; FILE REFERENCE: 01017/35966A
; CURRENT APPLICATION NUMBER: US/09/724,126A
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 60/187,211
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 18
; LENGTH: 5205
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (662)
; OTHER INFORMATION: n = A or T or G or C
; NAME/KEY: misc feature
; LOCATION: (668)
; OTHER INFORMATION: n = A or T or G or C
US-09-724-126A-18

Query Match 81.4%; Score 5136; DB 4; Length 5205;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 5189; Conservative 10; Mismatches 6; Indels 45; Gaps 1;
Qy 696 ATGCGGACGAGGAGGCTGAGAGGTACTGAGAGGATGGAATCAGCGCGAGGTATACCCAG 755
Db 1 ATGCGGACGAGGAGGCTGAGAGGTACTGAGAGGATGGAATCAGCGCGAGGTATACCCAG 60

[illegible]

QY 5136 TCTCAATATACAGTGGCTCCATTTGGGTGTGATATTTCTGGCTGTGATATTTGGGTCTCA 5195
Db 4396 TCTCAATATACAGTGGCTCCATTTGGGTGTGATATTTCTGGCTGTGATATTTGGGTCTCA 4455
QY 5196 CTGAAGATGGCATCACCCCTTATCTTCTGTGTGCTGATGTTTTCACACTATTACTT 5255
Db 4456 CTGAAGATGGCATCACCCCTTATCTTCTGTGTGCTGATGTTTTCACACTATTACTT 4515
QY 5256 GGGGTAACTCCGCTCGAGAACTGCATACCAATTTCTGCAGAGAGAGATGACATGCTC 5315
Db 4516 GGGGTAACTCCGCTCGAGAACTGCATACCAATTTCTGCAGAGAGAGATGACATGCTC 4575
QY 5316 TGTAGCTATCTATCTTTACCTACAAATTTGCTCTGCTCTTCCAGAAATATTGGATCT 5375
Db 4576 TGTAGCTATCTATCTTTACCTACAAATTTGCTCTGCTCTTCCAGAAATATTGGATCT 4635
QY 5376 GTAAGGCTCTGCTCCAGAGGTGTGTCAGATCTGCTTACTTAACTGTTTGAAGCA 5435
Db 4636 GTAAGGCTCTGCTCCAGAGGTGTGTCAGATCTGCTTACTTAACTGTTTGAAGCA 4695
QY 5436 AAAACACCGTGTGCTGAGTACCTAGAAAAAGAAATAGTTTGTATAGAGCTTCTGATGAC 5495
Db 4696 AAAACACCGTGTGAGTACCTAGAAAAAGAAATAGTTTGTATAGAGCTTCTGATGAC 4755
QY 5496 TATAGCTGCTCTGATCAATCAAGCTTCTCATTTCAAGTGTGCTGCTGATGATGAG 5555
Db 4756 TATAGCTGCTCTGATCAATCAAGCTTCTCATTTCAAGTGTGCTGCTGATGATGAG 4815
QY 5556 CGAAGCATCTGCTCTGCTCTGCTTCTGCTGGGCTATATGTTCTCAGAACATTTGC 5615
Db 4816 CGAAGCATCTGCTCTGCTCTGCTTCTGCTGGGCTATATGTTCTCAGAACATTTGC 4875
QY 5616 TGCCAGGAAATGTGAAGGGGAGAGGTGTGAGCTTGCATTTTTCAGCAGCTTCACTGT 5675
Db 4876 TGCCAGGAAATGTGAAGGGGAGAGGTGTGAGCTTGCATTTTTCAGCAGCTTCACTGT 4935
QY 5676 GGAGCGGAGTGTGATTTTCTTAAATCAGAGAAATCCGAGTGGTCTCGTTGAAGGT 5735
Db 4936 GGAGCGGAGTGTGATTTTCTTAAATCAGAGAAATCCGAGTGGTCTCGTTGAAGGT 4995
QY 5736 AAGCCAGAGCTGTGCTTATCCAGCTTCTTACTTGGATGATATGGAGAACAGACCT 5795
Db 4996 AAGCCAGAGCTGTGCTTATCCAGCTTCTTACTTGGATGATATGGAGAACAGACCT 5055
QY 5796 GGCCTGAAGAGGGGCAACCCCTTCTATTTATCTGCTGAGCGGTATCGGAAGCTCCATTG 5855
Db 5056 GGCCTGAAGAGGGGCAACCCCTTCTATTTATCTGCTGAGCGGTATCGGAAGCTCCATTG 5115
QY 5856 GTCTGGCAACAACTGCTATATAGAGAGATTGCTAGGAGCAAGAGACTAATCAGATG 5915
Db 5116 GTCTGGCAACAACTGCTATATAGAGAGATTGCTAGGAGCAAGAGACTAATCAGATG 5175
QY 5916 TTATTTGATTTCACTGCGAGTACTGTGA 5945
Db 5176 TTATTTGATTTCACTGCGAGTACTGTGA 5205

RESULT 3

US-08-982-956-1
; Sequence 1, Application US/08982956
; Patent No. 5861312
; GENERAL INFORMATION:
; APPLICANT: Varshevsky, Alexander
; APPLICANT: Kwon, Yong Tae
; TITLE OF INVENTION: NUCLEIC ACID ENCODING MAMMALIAN UBRI
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kevin M. Farrell, P. C.
; STREET: P. O. Box 999
; CITY: York Harbor
; STATE: ME
; COUNTRY: US

ZIP: 03911
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/982,956
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Farrell, Kevin M.
; REGISTRATION NUMBER: 35,505
; REFERENCE/DOCKET NUMBER: CIT-2001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (207) 363-0558
; TELEFAX: (207) 363-0528
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6395 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 115..5385
US-08-982-956-1

Query Match 70.3%; Score 4435.4; DB 2; Length 6395;

Best Local Similarity 86.3%; Pred. No. 0;

Matches 4997; Conservative 0; Mismatches 721; Indels 69; Gaps 6;

QY 590 TTCAGGGGCGCTCGTAAAGTGTGCTCCCTGTCTC-TCCGACCGGCCACAGGTTTCCGCT 648
Db 8 TTCAGGGGCGCTCGTAAAGTGTGCTCCCTGTGCGGTGCGGCGGCCACAGGTTTCCGCT 67
QY 649 TGCCTCTTGGCGGGGTGGGCAACTGCAAGGCGTCAAGTTTCCCTCAAGATGGCGGACGAGG 708
Db 68 AGCTGGCGCGCGGGGTGGGCAACTGCGGGCGTTCGTTTCCCTTAAGATGGCGGACGAGG 127
QY 709 AGGCTGGAGGTACTGAGAGGATGGAATCAGCGGGAGTTACCCAGACCCCTCAGCGTC 768
Db 128 AGATGGACGGCGCGGAGAGGATGGACGTCAGCCCGAGCCTCCCTTGGCCCGCAGCGGC 187
QY 769 TGGCATCTTGTGTGGATCAGCAAGTTGATTTTATCTGCTTTTCTGTCATCATTTGGCAC 828
Db 188 CGGCATCTGTTGGGATCAGCAAGTTGATTTTCTATCTGCTTTCTTACATCATTTGGCAC 247
QY 829 AATTGGTCCCAAAATTTTACTTTCTGAAATGGACCCAGACTTGGAAAGCAGGAGGAAA 888
Db 248 AATTAGTGCCAGAAATTTTATTTTCTGAGATGGACCCAGATTTGGAAAGCAAGAGAGA 307
QY 889 GTGTACAAATGTCAATTTCACTCCATCTGGAATGGTACTTATTTGGAGAGATCCAGATA 948
Db 308 GTGTACAGATGTCAATTTCACTCCCTTTGGAGTGGTACTTATTTGGAGAGATCCGATA 367
QY 949 TTTGCTTAGAGAAATTCAGACACAGTGGACATTTTCACTTTTGGGAGGCTTTTCAAAA 1008
Db 368 TTTGCTTAGAGAAATTCAGACACAGTGGAGGCTTCCAGTTGTGTGGAGGCTTTTCAAAA 427
QY 1009 GTGGAGAGACAACTTATTTCTTGGAGGATTTGTCAATTTGATTCACCAATGTGTACTCTGA 1068
Db 428 GTGGAGAGACAACTTATTTCTTGTAGGATTTGTCAATTTGATTCACCAATGTGTGTCTCTGA 487
QY 1069 TGGATCTTCCAGGACAGTGTTCATAAATCATGTTTACAGATGATGATCTTACTCTG 1128
Db 488 TGGATCTTCCAAAGTAGTGTTCATAAAAACCATCGTTTACAGATGATGATCTTACTCTG 547
QY 1129 GAGGAGGCTTCTGTGACTGTGGAGACACAGAGGATGGAACCTGGCCCTTTTGTGTAA 1188
Db 548 GAGGAGGCTTCTGTGACTGTGGAGACACAGAGGCTGGAAAACTGGCCCTTTTGTGTGG 607

Db 2768 ATGTTATGATATACATCTCTCAGGACCATCTTTTGAGCGGCGAGTGACACGGAGTCTAATC 2827
Qy 3409 TGTGACCGAAGGGATGCTCCAAATGGCTTTTCATATTTCTGCGATGCGTGTATCTAGAG 3468
Db 2828 TGTGACAGAAAGGGATGCTGACAGATGGCGTTCCATATATTTGGCATCTGGGCTTGTCTGAAG 2887
Qy 3469 AGAAGCAACAGCTTCCAAAGAGCTCTGAAGAGAGAGTAACATTTGACATTTTATCATNAGG 3528
Db 2888 AGAAGCAGCAGCTTCGAAAGCTCTGAAGAGAGAGTGGCTTTTGACATTTTACCATNAG 2947
Qy 3529 CTTCAAGATTGGGAAGTTTCAGCCAT- - - - -GAATATACAAATGCTTTTGGAAAAAC 3579
Db 2948 CTTCAAGATTGGGAAGTTTCAGCCATGAATGCTCAGAAATATACAAATGCTCTTGGAAAGAC 3007
Qy 3580 TCAAGGAATTCCTCCAGTTAGAGGCCAGAGGACATGATACGTGGATATCTTCAGATGT 3639
Db 3008 TCAAGGAATTCCTCCAAATTAGAGGCCAGAGGACATGATACATGATATCTCCAGATGT 3067
Qy 3640 TTGACACAGTGAAGGATTAAGAGAAAATCTTGTTTAATTTGTAGACACCATCAGCAT 3699
Db 3068 TTGACACAGTGAAGGATTAAGAGAAAATCTTGTTTAGTTTGTGGCAACCATCTCAGGAC 3127
Qy 3700 CGGAATCTATTAAGGAATGATGAGATTAATCTCATGATATAAGAAAAAGACGAAAGAAAGNA 3759
Db 3128 TGGAGTGCAATTAAGAGTGAAGGATTAATCTCATGATATAAGAAAAAGGACAGAAAGGAGNA 3187
Qy 3760 AAGCTGAAGCTGTAGGCTATCATGCCAGAGAGATCATGGCTCAGATGTCTGCTTACAGA 3819
Db 3188 AAGCTGAGCGCGCTAGGCTTTCATGCCAGAGAGATCATGCCAGAGATGTCTGCTTACAGA 3247
Qy 3820 AAAATCTCATTAAGAACTATAAATCATGATGATGACATACATCAGAAATGCTCGGGAAG 3879
Db 3248 AAAATCTCATTAAGAACTATAAATCATGATGATGACATACATCAGAAATGCTCGGGAAG 3307
Qy 3880 AAGATTCATTAAGGAAGAGAGACACCCAGAGAGTCAAGTCACTCTAGAAATTTGCTT 3939
Db 3308 AAGACTCAATTAAGGAAGAGAGACCTCAGCAGTCAAGTGAAGGCTCTAGAAATTTGCTC 3367
Qy 3940 TGGGTCTTAAGCGGGGCTCATCTGTTACTGAAAAAGAGGCTGCTGAGTGCAATCTTTGCC 3999
Db 3368 TGGGCGCTTAAGCGGGGCGGCTGTTTACCGAAAAAGAGGCTGCTGAGTGCAATCTCTGCC 3427
Qy 4000 AAGAGACAGAGGTGAAATAGAAATATGCAATGATGATATTCGCGCTGTGTCCAGA 4059
Db 3428 AAGAAGAACAAAGAGGTAAATCTAGAAAAATATGCCATGTTTGTACGATGTGTGCGAGA 3487
Qy 4060 AATCTACTGCTTAACCCAGACAGGGGAAACCCATAGAACTCTCAGGAGAACCCCTAG 4119
Db 3488 AATCACCGCCCTTAACCCAGACAGAGGAAAGCTGTGACCACTTGAAGGAAACACTGG 3547
Qy 4120 ACCCACTTTTCATGGATCCAGACTTTGGCATATGGAACCTTATACAGGAAGCTGTGTGATG 4179
Db 3548 ACCCTCTTTTCATGGATCCAGACTTTGGCATATGGAACCTTATACAGGAAGCTGTGTGATG 3607
Qy 4180 TAATGACAGAGTGTGTGGCAGAGATTTTGAAGCTGTACAGTGAAGCTCTCAGGAGC 4239
Db 3608 TAATGACAGAGTGTGTGGCAGAGATTTTGAAGCTGTGAGCTGTGAGCTGTGAGCAGC 3667
Qy 4240 GCATTCATGTTGACCTTTTTCACCTTGGAAAGTGGAGATATCTTTGCCCTCTTTGCAAT 4299
Db 3668 GCATTCAGTAGACCTTTTTCACCTTGGAGAGCGGAGTACTATAGCCGCTCTGCAAGT 3727
Qy 4300 CTCTGTGCAATCTGTGATCCCAATTAATCTTTTTCGCAACCTCAAAAGATAACAGTGA 4359
Db 3728 CTCTGTGCAACACTGTATCCCATCATCCCTTTTGGCAGCGCAGAGATCAACAGTGA 3787
Qy 4360 ATGACAGTGTCTGTCTCAATCTTTTGAACCTTGGCAGCGGTGATACAGACTGTTCTGGCCA 4419
Db 3788 ATGGGAGGCTCTTGTCTCAATCTTTTGAACCTTGGGCGCGGTGATACAGACTGTCTTGCCA 3847
Qy 4420 GAATATCAGGTTAATAATAGACATGCTAAAGGAGAAAC- - -CAATTCCTATTTTCT 4476
Db 3848 GAATATCGGGTTAATAATAGACATGCTAAAGGAGAAAGCCCGCAGCAGTTCCTGTCTGT 3907

Qy 4477 TTAATCAAGGAATGGGAGATTTCTACTTTGGAGTTTCCATTTCCATCTCTGAGTTTGGCGTTG 4536
Db 3908 TTAATCAAGGAATGGGGGATTTCAACTTTTGGAGTTTTCATTTCCATCTCTGAGTTTGGAGTTT 3967
Qy 4537 AGTCTTCGATTTAAATATTTCAAAATAGCATCAAGGAAATGGTTATTCTCTTTTGGCCACAA 4596
Db 3968 AGTCTTCGTTGAAATATTTCAAAATAGTATCAAGGAAATGGTCTATCTCTTTCGCCACAA 4027
Qy 4597 TTTATAGAAATTTGAAATGCACTGCACTGATGAAGGGAATCTCTGAGTCCCCATGCTGA 4656
Db 4028 TTTACAGAAATTTGGCTGAAAGTGCCTCTGATGAATAGACCCAGAGTGCCTCATGATGA 4087
Qy 4657 CTTGGAGCACCTGCGCTTTTCACTATCCAGGCAATTTGAAATCTATTGGGAGATGAGGAA 4716
Db 4088 CTTGGAGCACCTGCGCTTTTCACTATCCAGGCAATTTGAAATCTATTGGGAGATGAGGAA 4147
Qy 4717 AACCTCTGTTTGGAGCACTTCAAAATAGGAGCAATATGGTCTGAAAGCAATTAATGCAAT 4776
Db 4148 AACCTCTATTTTGGAGCACTTCAAAATAGAGCAATAGCGGTCTGAAGCGGCTTAATGCAAT 4207
Qy 4777 TTGCAGTTGCAAGAGGATTAACCTGTCTCAGGCTCTGATACAGAAACATCTGTTCTGTC 4836
Db 4208 TTGCAGTTGCAAGAGGCTTAACCTGTCTCAGGCTCTGATACAGAAACATCTGTTCTGTC 4267
Qy 4837 TTCTATCAGTTTGTCTTCTTAACATAAATCAGAAAGATACACATGCTCTTCTGCTATAG 4896
Db 4268 TCCTGTCAAGTTATTCTTCTTAACCTGCAATCAGAAAAATACACAGGCTCTTCTGCTG 4327
Qy 4897 ATCTGTTTCAATGTTTGGTGGTGTCTGTTGATGATTTCCCATCTGTTATTTGGGATGAC 4956
Db 4328 ATCTGTTTCAATGTTTGGTGGTGTCTGTTGATGATTTCCCATCTGTTATTTGGGATGAC 4387
Qy 4957 CTGTTGATCTGAGCCTTTCTTCAATGATTTCTTCTTAAACCACTTTTATCTTCTTCAAT 5016
Db 4388 CCGTGGATCTGAGCGCTGCGCACTTATGTTCTTCAATATAACCACTTCTTCTTCTTCAAT 4447
Qy 5017 TGATCACATGCGCACATGCTTTCAGATATCTTCAAGTAGACA- - - - -CAG 5064
Db 4448 TGATCACATGCGCGCACATGCTTTCAGATATCTTCAAGTAGACA- - - - -CAG 4507
Qy 5065 GCCTACCCCTTGTCTCAGGTTCAAGAACAGTGAAGAGGCTCATTTCCGCACTTCTTCTTCT 5124
Db 4508 GCGCGCGCTTGTCTGAGGGTGAAGAGGATAGTGAAGGCTCGCTGTGTCATCTGCTTTCT 4567
Qy 5125 TTGCAGAAATTTCTCAATATACAGTGGCTCCATTTGGTGTGATATTTCTGCTGGCTGAT 5184
Db 4568 TTGTGGAAGTGTGCGACGACACAGACGCGCTCTACTGCGGTGGGTGCTCCCGCTGGTACC 4627
Qy 5185 TGTGGGTCTCACTGAAGAAATGGCATCACCCCTTATCTTGTGCTGTGCTGCTGCTTTTTC 5244
Db 4628 TGTGGCTCTCCCTGAGGAAAGGCAATCACCCCTTACCTCCGCTGCTGCTGCTGCTTTTTC 4687
Qy 5245 ACTATTTACTTGGGTTAACTCCGCTGAGGAACTGCAATCCAAATTTCTGCAATTTCTGAGAG 5304
Db 4688 ACTATTTACTTGGGTTAGCTCCGCTGAGAACTGTTTGGCAATTTCTGCTGAAGAGAAAT 4747
Qy 5305 ACAGTGCACTCTGTAGCTATCTATCTTTTACCTAATAATTTGTTCTGCTGCTTTCAGGAAT 5364
Db 4748 TCAGTGCACTCTGTAGCTATCTATCTTTTACCAATAATTTGTTCTGCTGCTTTCAGGAAT 4807
Qy 5365 ATTGGGATCTGTAAAGCCCTTGTCTCAGAGGTTGGTGTGAGATCTCTGCTTACTTAACT 5424
Db 4808 ATTGGGATCTGTAAAGCCCTTACTTACAGAGGTTGGTGTGAGATCTCTGCTTACTTAACT 4867
Qy 5425 GTTTGAAGCAAAAAACACCGTGTGAGTACCTTAGAAAAAGAAATAGTTTGTATAGAGC 5484
Db 4868 CTTTGAAGCAAAAAAGTGTGTGCTGAGTACCTTAGAAAAAGAAATAGTTTGTATAGAGC 4927
Qy 5485 TTCTGTATGATATAGCTGCTCTGATCAAGCTTCTCAATTTCAAGTGGCCACCGTCTG 5544
Db 4928 TTCTGTAGGACTACAGCTGTCTTCTAATCAGGCTTCTTCAATTTAGGTGTCCAGGTTCTG 4987

```
QY 5545 CAGATGATGAGCGAAGCATCTGTCCTCTGCTTTCTGTGGGCTATATCTATGTTCTC 5604
Db |||||
QY 4988 CAGATGATGAGCGAAGCATCTGTCCTCTGCTTTCTGTGGGCTATCTGTGTTCTC 5047
Db |||||
QY 5605 AGAACATTTGCTGCCAGGAAATTTGAAACGGGGAAGAGTTTGGAGCTTTGCAATTTTTCACG 5664
Db |||||
QY 5048 AGAACATCTGTTGCCAAGAAATAGTAATGGAGAGAGTTTGGAGCGTGGCTTTTTCATG 5107
Db |||||
QY 5665 CACTTCACTGTGGAGCCGAGTCTGCTATTTTCTTAAATCAGAGAAATGCCAGAGTGGTCC 5724
Db |||||
QY 5108 CGCTTCAATGTTGCTGAGAGCTGCTGCTATTTTCTTAAATCAGAGAAATGCCAGGTTGGTCC 5167
Db |||||
QY 5725 TGGTTGAAGTAAAGCCAGAGGCTGTGCTATCTCAGCTCTTACTTTGATGAATATGAG 5784
Db |||||
QY 5168 TGGTGAAGGAAAGCCAGAGGCTGTGCTATCTCAGGCTTCTTACTTTGATGAATATGAG 5227
Db |||||
QY 5785 AAACAGACCCCTGGCTGGAAGGGGCAACCCCTTCAATTTCTGCTGAGCGGTATCGGA 5844
Db |||||
QY 5228 AAACAGACCCAGGCTAAGAGAGGAAACCCACTTCAATTTCTGCGGAGCGGTATCGGA 5287
Db |||||
QY 5845 AGCTCCATTTGCTGGCAACAACTGATATAGAGAGATTTCTAGGAGCGCAAGAGA 5904
Db |||||
QY 5288 AGCTGCATTTGCTGCGCAACAGCACTGCAATATATAGAGAGATTTCTGCGAGCCAGGAGA 5347
Db |||||
QY 5905 CTAATCAGATGTTATTTGGAATCAACTGCGAGTACTGTGAGCTCCAACCTCTGCTCAAG 5964
Db |||||
QY 5348 CTAATCAGATGCTATTTGGAATTAAGTGGAGTACTGTGAGCTTCAAGTCTGCTCAAG 5407
Db |||||
QY 5965 ACAATCACAATAGCAGCATAGTAAGAGGCTGATTTCAAAATTTATGGAAGAACTTTCTGAGG 6024
Db |||||
QY 5408 ACAATCAGTGCATCAATAATAAGACTGATCTAAATTTCTAGAGAACTTTCTGAGG 5467
Db |||||
QY 6025 GCTGGGAAGTATTTGAGGGGCTTTTGTCTCAATGTCAGGTTTCACTTACATCAATAAAT 6084
Db |||||
QY 5468 ACGGGGGAAGTATTTGAGGGGCTTTTGTATCCATGTCAGATTTCAACACATTAATAAAT 5527
Db |||||
QY 6085 ATTTCTTAATGAGTATTTGCTTCAATTTAGCAACATATGCTTCAACAGAAAAA-AGGAC 6143
Db |||||
QY 5528 ATTTCTTAATGGAATTTGCTTCAATTTATCAACATAGCTTCAAGGGAAGAAAAACAGAC 5587
Db |||||
QY 6144 ATAGATCAATCTGT-----TTT 6160
Db |||||
QY 5588 ATAGATTAATGTTTATGTTCTAGAACACTAAAGAAATGCTTTGTTCAATCAAGTCTAT 5647
Db |||||
QY 6161 ATGCTAGTATTTCCAGGAATTTATCCCTTCAATAATTTGCTCAATTTCAATTTATTT 6220
Db |||||
QY 5648 TTCTGCTAATATTTCCAGAAAACTCTTTCCCTTCAATACTGCTTAGTTCAATTTCAAT 5707
Db |||||
QY 6221 CATCCACTTGGTAGATGAAGTCACTCAACAGTTGTAGACATTTTATGTTGGTTAAC 6280
Db |||||
QY 5708 CACCACCTGGTAAATGAGGTACATTAAGCAATTTGTGACATTTTCTCCATCTGGCTAAC 5767
Db |||||
QY 6281 TCTTCTGCAATTTTGTATTTGGTGT 6307
Db |||||
QY 5768 ATCTCTGCACCTTTGATTTGGTGT 5794
Db |||||
```

RESULT 4

```
US-09-228-317-1
; Sequence 1, Application US/09228317
; Patent No. 6159732
; GENERAL INFORMATION:
; APPLICANT: Varshavsky, Alexander
; APPLICANT: Kwon, Yong Tae
; TITLE OF INVENTION: NUCLEIC ACID ENCODING MAMMALIAN UBRI
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kevin M. Farrell, P.C.
; STREET: P.O. Box 999
; CITY: York Harbor
; STATE: ME
; COUNTRY: US
; ZIP: 03911
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/228,317
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Farrell, Kevin M.
; REGISTRATION NUMBER: 35,505
; REFERENCE/DOCKET NUMBER: CIT-2001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (207) 363-0558
; TELEFAX: (207) 363-0528
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6395 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 115..5385
; US-09-228-317-1
```

Query Match 70.3%; Score 4435.4; DB 3; Length 6395;

Best Local Similarity 86.3%; Pred. No. 0;

Matches 4997; Conservative 0; Mismatches 721; Indels 69; Gaps 6;

```
QY 590 TTCAAGGGCCGCTCTAAAGTGTGTCCTCTGCTCTC-TCCGACCCGCCACACAGTTTCGCT 648
Db 8 TTCAAGGGCCGCTCTAAAGTGTGTCCTCTGCTCTCAGCGTCGGGCCCGGCCACAGTTTCGCT 67
QY 649 TGCCTCTGCGCGCGGCTCGGCACTGCAGAGGCTCAGTTTCCCTCAAGATGCGGACGAGG 708
Db |||||
QY 68 AGCTGGCGCGCGGGGCTCGGNACTGGGGGCTTGGTTTCCCTTAAGATGCGGACGAG 127
Db |||||
QY 709 AGGCTGGAGGTACTGAGAGGATGGAATCAGCGCGGAGTTACCCAGACCCTTCAGCGTC 768
Db |||||
QY 128 AGATGGAGCGCGCGAGAGGATGACGTCAGCCCGGAGCTCCCTCGGCCCGCGAGCGGC 187
Db |||||
QY 769 TGGCATCTTGTGGGATCAGCAAGTTGATTTTATATCTGCTTTCTTGATCATTTGGCAC 828
Db |||||
QY 188 CGGCATCTGTGGGATCAGCAAGTTGATTTCTATCTATGCTTTCTTACATCATTTGGCAC 247
Db |||||
QY 829 AATTGGTCCAGAAATTTACTTTCTGAAATGGACCCAGACTTGGMAAAGCAGAGGAAA 888
Db |||||
QY 248 AATTAGTCCAGAAATTTATTTTCTGAGATGGACCCAGATTTGGMAAAGCAGAGAGA 307
Db |||||
QY 889 GTGTACAAATGTCAAATTTCACTCCACTGGAAATGGTACTTATTTGGAGAGATCCAGATA 948
Db |||||
QY 308 GTGTACAGATGTCAAATCTCACTCTCTTTGGAGTGGTACTTATTTGGAGAGATCCGATA 367
Db |||||
QY 949 TTTGCTTAGAATAATTTGAAGCAGATGAGCAATTTGCTGTTGGGAGGTTTTCAAA 1008
Db |||||
QY 368 TTTGCTTAGAATAATTTAAAACACAGATGAGCGTTTCCAGTGTGTGGGAGGTTTTCAAA 427
Db |||||
QY 1009 GTGGAGAGACAACCTTATTTTTCAGGGATTTGCAATTTGATCCAACTGCTACTCTGTA 1068
Db |||||
QY 428 GTGGAGAAACCAATATTTCTGTAGGATTTGCAATTTGATCCAACTGCTACTCTGTA 487
Db |||||
QY 1069 TGGACTGCTTCCAGGACAGTGTTCATAAAAATCATCGTTTACAAGATGCATCTTACTG 1128
Db |||||
QY 488 TGGACTGCTTCCAAAGTAGTGTTCATAAAAACCATCGTTTACAAGATGCATCTTACTG 547
Db |||||
QY 1129 GAGGAGGTTTCTGTGACTGTGGAGACACAGAGGATGGAATACTGGCCCTTTTGTGTA 1188
Db |||||
QY 548 GAGGGGCTTCTGTGACTGTGGAGACACAGAGCGTGGAAAACCTGGCCCTTTTGTGTTGG 607
Db |||||
QY 1189 ATCATGAACCTGGAAAGACAGGTACTATATAAAGAGAAATTCACGCTGTCCGTTGATGAAG 1248
Db |||||
```

[illegible]

Db	1688	ACATTGAAGTGGACCCCTGACTGGGAGGCTGCCATCGCTATACAGATGCAACTAAAGAATA	1741
Qy	2329	TTTTTACTCATGTTCCAAAGAGTGGTGTGCTTGATGAAGAACCTCTTACTTGTGCGCTTATA	2388
Db	1748	TTTTTGCTCATGTTCCAGAGTGGTGTGCTTGATGAGAGATCTCTTACTGGTGGCTTATA	1807
Qy	2389	AAGAAATGTCAAAAGCTGTGATGAGGTGCAGTACCAGTTTTCATATCTAGTGTAGCAAGACAG	2448
Db	1808	AAGAAATGTCAAAAGCTGTAAATGAGGTGCAGTACAAATTTTCATGTCCAGTACCAGACAG	1867
Qy	2449	TAGTACAAATCGTGTGGACATAGTTTGGAAACAAAGTCCTACAGAGTATCTGAGGATCTTGG	2508
Db	1868	TAGTGCAAATTTGTGGGGTCATAGTCTGAAACCAAAATCCTCAAAAGTGTCTGAGGACCTTG	1927
Qy	2509	TAAGCATACATCTGCCACTCTAGGACCCCTGCTGTGCTCTTCAATGTACGTGTTTAAGCAGGC	2568
Db	1928	TAAGCATACACCTGCCACTCTTAGAACACTTGTCTGTCTTCATGTACGTGTTTAAGCAGAC	1987
Qy	2569	TGGTGCTGTTTCAAGACTCGATGAATTTGTGTCTTTTGAAGACTTTTCAAGTATAGGTATC	2628
Db	1988	TAGTGTCTTAATTTCAAGACTCGATGAATTTGTGCTTTTGAAGCTTTTCAAGTATAGGTATC	2047
Qy	2629	TAGTGGAAATATCCTTTTACGTGTCTGGTGTGCTTGGCCAGGTTGTCTGAGATGTGGC	2688
Db	2048	TGSGGAGTACCCCGCTGCGCTGCTGCTCTGCTGTGCTCAGGTGTGTCTGAGATGTGGC	2107
Qy	2689	GAAGAAATGCACTGTCTCTTATTAGCCAGGTGTTTATTACCAAGATGTTTAAGTGCAGAG	2748
Db	2108	GAAGAAACGGGCTCTCACTCATCAGCCAGGTTTCTATTATCAAGATGTTTAATGCAGGG	2167
Qy	2749	AAGAAATGTATGATAAAGATATCATCATGCTTCAGATGTGTGTGCATCTTTAATGGATCCCA	2808
Db	2168	AGGAAATGTACGATAAAGATATCATCATGCTTCAGATTTGGAGCATCTATAATGGAATCCCA	2227
Qy	2809	ATAGTTCCTTTTACTGTTACTTCTCAGAGGTATGAACTTTCGGAGGCTTTTAAACAAGCCA	2868
Db	2228	ACAAGTTCCTTTTACTGTTACTTCTCAGAGATATGAACCTTTACTGATGCTTTTAAACAAGCCA	2287
Qy	2869	TATCTCAAAAGACCCAGGATTTGATTTAAACAAATAATAACATACATATAGAGAATAATGCTTC	2928
Db	2288	TATCCAAAGAACCAGGATTTGATTTAAACNGTATATATACATTTAATAGAAGAAATGCTTC	2347
Qy	2929	AGTTCCTCATATATTGTGGGTGAGCGTTATGTACTCTGGAGTGGAAATGTGACCAAG	2988
Db	2348	AGTTCCTCATATATTGTGGGAGAACGTTATGTACTCTGGAGTGGGAAATGTTACCAGAG	2407
Qy	2989	AAGAGTCACAATGAGAGAAATCAITTCATCTTGTCTTGTTCATTGAAACCCATGCCACACAGTG	3048
Db	2408	AGGAGGTTAATTAAGAGAGATTTACTCCTCTCTTTGCAATTTGAGCCCTATGCCACACAGTG	2467
Qy	3049	CCATTGCCAAAAATTTTACCTCGAGAATGAAATAATGAAACTGGCTTTAGAGAATGTCTATAA	3108
Db	2468	CCATGCCAGAAACCTTACCTGAGAACGAAATAATGAAACTGGCTTTAGAGAATGTCTATAA	2527
Qy	3109	ACAAAGTGGCCACATTTAAGAAACCGAGGTGTATPCAGGCCATGGAGTTTATGAACTAAAAAG	3168
Db	2528	ACAAAGTGGCCACATTTAAGAAACCGAGGTGTGTGCGGCCCATGAGTTTATGAAATGAAAG	2587
Qy	3169	ATGAATCACTGAAGACTTCATATGTACTCTTTTATCATTTACTTCCAAAAACCCAGCATAGCA	3228
Db	2588	ATGAATCACTGAAGACTTCATATGTGTACTTTTACCATTTATTTCTAAAAACACAGCATAGCA	2647
Qy	3229	AGGCTGAACATATGCAAGAAAGAGGAAAAACAAGAAAAACAAGATGAAGCATTGCCCG	3288
Db	2648	AGGCTGAACATATGCAAGAAAGAGGAAAAACAAGAAATAAAGATGAAGCATTGCCCG	2707
Qy	3289	CACCACCACTCTCTGAATTTCTGCGCTCTTTTCAGCAAAAGTGAATTAACCTTCTCAACTGTG	3348
Db	2708	CGCCACCTCTCCAGAGTCTTCTGCGCTCTTTTCAGCAAAAGTGAATTAACCTTCTCAACTGTG	2767
Qy	3349	ATATCATGATGTACATTTCTCAGGACCGTATTTTGGCGGGCAATAGACACAGATTTCTAATCT	3408
Db	2768	ATGTTATGATATATCATCTCTCAGAGCACTCTTTTGGCGGGCAGTGGACAGGAGTCTTAATCT	2827

Qy	3409	TGTGGACCGAAGGGATGCTCCAAATGGCTTTTCATATTTCTGGCAATTTGGGTTTACTAGAAG	3468
Db	2828	TGTGGACAGAAGGGATGCTGCAGATGGCTTCCATATATTTGGCACTGGCTTGTGTGAAG	2887
Qy	3469	AGAAGCAACAGCTTCAAAAAGCTCTCGAAGAGAAGTAACAATTTGACTTTTATCATNAGG	3528
Db	2888	AGAAGCAGCAGCTTCAGANAAGCTCTCGAAGGAAGTGGCTTTTGTACTTTTACCATTAAG	2947
Qy	3529	CTTCAAGATTGGGAAGTTTCAGCCAT-----GAATATACAAATGCTTTTGGAAAAAC	3579
Db	2948	CTTCAAGATTGGGAAGTTTCAGCCATGAATGCTCAGAATATACAAATGCTCTTGGNAAGAC	3007
Qy	3580	TCAAGGAATTTCCCACTTGAAGGCCCAGAAGGACATGATAACGTGGATACTTCCAGATGT	3639
Db	3008	TCAAGGAATCCCCCAATTAGAAGGCCAGAAGGACATGATAACATGATGATCTCCAGATGT	3067
Qy	3640	TTGACACAGTGAGAGCGAATTAAGAGAAAAATCTTGTTTAAATTTGTAGCAACACACATCAGAT	3699
Db	3068	TTGACACAGTGAGAGCGAATTAAGAGAAAAATCTTGTTTAAATTTGTGGCAACACATCTCAGGAC	3127
Qy	3700	CGGAATCTTATTAAGAAATGATGAGATTACTCATGATAAAGAAAAAGACAGAAACGAAAAAGAA	3759
Db	3128	TGGAGTGCATTTAAGGTGAGGAGATTACTCATGATAAAGAAAGACGACAGACGAGNAGAA	3187
Qy	3760	AAGCTGAAGCTGTAGGCTCATCGCCAGAAGATCATGGCTTCAGATGTCTGCCCTTACAGA	3819
Db	3188	AAGCTGAGGCGCTAGGCTTCATCGCCAGAAGATCATGGGCCAGATGTCTGCCCTTACAGA	3247
Qy	3820	AAAACTTCAATTGAAACTCATAACTCATGTATGCAATATCATCAGAAATGCCCTGGGAAG	3879
Db	3248	AAAACTTCAATTGAAACCCCAAACTCATGTATGATAATACTCGTCAAGAATTAACAGGGAAGG	3307
Qy	3880	AAGATTCCATTTATGGAGGAGAGACACCCAGCAGTCAGTGACTACTCTAGAAATTCGCTT	3939
Db	3308	AAGACTCCATTTATGGAGGAGAGACACCTCAGCAGTCAGTGAAGCTCTTAGAATTCGCTC	3367
Qy	3940	TGGTCTCTAAACGGGGTCCATCTGTTTACTGAAAAGAGGTGCTGACGTGCATCTTTGGCC	3999
Db	3368	TGGGCCCTTAAACGGGGCCGGCTGTTTACCGAAAAGAGGTGCTGACGTGCATCTCTGGCC	3427
Qy	4000	AAGAAGAACAGGAGGTGAAATAGAAATATATGCCATGTGTTATTCGGCTGTGTCCAGA	4059
Db	3428	AAGAAGAACAGAGGTAAACTAGAAATATATGCCATGTGTTATTCAGCATGTGTGCAGA	3487
Qy	4060	AATCTACTGCCCTTAACCCAGCACAGGGGAAAAACCATAGAACTCTCAGGAGAAGCCCTAG	4119
Db	3488	AATCCAACGCCCTTAACCCAGCACAGAGGGAAGCCGTGTGACCACTTAGGGGNAACACTGG	3547
Qy	4120	ACCCACTTTTCAATGGATCCAGACTTGGCATATGNAACCTTATACAGGAAGCTGTGGTCATG	4179
Db	3548	ACCCCTCTTTTCAATGGATCCAGACTTGGCACAATGGAACCTTATACAGGAAGCTGTGGTCATG	3607
Qy	4180	TAATGCACGCACTGTCTGGCAGAAATATTTTGAAGCTGTACAGCTCAGACTCTCAGCAGC	4239
Db	3608	TAATGCATGCACTGTCTGGCAGAAATATTTTGAAGCTGTGCACTGTGCACTGTGCGCAGC	3667
Qy	4240	GCATTCAGCTTTTGTGACTTTTGAATTGAAAAGTGGAGAATATCTTTGGCCCTCTTTGCAAT	4299
Db	3668	GCATTCAGCTAGACCTGTGTTGACCTGGAGAGCGCGAGTAGCTATGCCCCGCTCTGCAAGT	3727
Qy	4300	CTCTGTGCATACTGTGATCCCAATTAATCTTTTGGCAACCTCRAAAGATTAACAGTACAGA	4359
Db	3728	CTCTGTGCACACTGTCAATCCCAATCATCCCTTTTGCAGCCCGCAGAAATTAACAGTACAGA	3787
Qy	4360	ATGCAGATGCTCTTGTCTCAACTTTTGTGACCTGGCAGCGTGGATACAGACTGTCTCGGCCA	4419
Db	3788	ATGCGAGGCTCTTGTCTCACTTTTGACCTTGGCCCGGTGGATACAGACTGTCTCTGCCA	3847
Qy	4420	GAATATCAAGTTTATAATATAAGACATGCTTAAAGGAGAAAAAC-----CAATTCCTATTTTCT	4476
Db	3848	GAATATCGGGTTTATAATATAAGCATGTCTAAGAGAAAGCCCCCAGCAGTTCCTGTCTGT	3907

QY	4477	TTAATCAAGGAATGGGAGATTCTACTTTTGAGTTCATTTCCATCCTGAGTTTTGGCGTTG	4536
DB	3908	TTAATCAAGGAATGGGGGATTCAACTTTTGAGTTTTCAATCCATCCTGAGTTTTGGAGTTT	3967
QY	4537	AGTCTTCGATTTAAATAATTTCAAATAGCATCAAGGAAATGGTTATTTCTTTTGCACACA	4596
DB	3968	AGTCTTCGGTGAANAATTTCAAATAGTATCAAGGAAATGGTCATTTCTTCGCCACAACA	4027
QY	4597	TTTATAGAATTTGGANTTGAATGTCCACTCATGAAGGAGTCTTCGAGTCCCANTGCTGA	4656
DB	4028	TTTACAGAAATTTGGCTTGAATGTCCCTCTGATGAACTAGACCACGAGTGCCCATGATGA	4087
QY	4657	CCTGGAGCACCTGGCGTTTCACTATCAGGCAATTTGAAAATCTATTATTTGGGAGATGAAGAA	4716
DB	4088	CCTGGAGCACGTGTGCGTTTCACTATCAGGCAATTCGAAAACTGTTTGGGAGATGAAGAA	4147
QY	4717	AACCTCTGTTTGGAGCATTTCAAATATGAGGAGCATTAATGCTCTGAAAGCATTTAATGCAGT	4776
DB	4148	AACCTCTATTATTTGGAGCATTTCAAATATAGACAGATATGCGGTCTGAAAGGCGCTAATGCA	4207
QY	4777	TTGCAGTTTGACACAGAGATTTACCTGTCTCAGGTCTCTGATACAGAAACATCTGTTCTGTC	4836
DB	4208	TTGCAGTTTGACACAGAGGCTTACCTGTCCCTCAGGTCCTGATACACAAACATCTGGCTCGC	4267
QY	4837	TTCTATCAGTTGTTTCTTCTTAACATAAAATCAGAAGATACCAATGCGCTTCTGTCTATAG	4896
DB	4268	TCCTGTTCAGTTATTCTTCTTAACTGAGTATTTCTCTATAAACCACTTTATCTCTTCCATT	4327
QY	4897	ATCTGTTTTCAGTTTGTGGTGGGTCTGTGTAGCATTTCCCATCTGTGTATTTGGAGTACC	4956
DB	4328	ATCTCTTCCAATTTCTGGTGGGCGCAGTCTTAGCGTTTCCCATCTCTGTATTTGGGATGACA	4387
QY	4957	CTGTGTGATCTGCAGCGCTTCTTTCAGTTAGTTTCTCTCTATAAACCACTTTATCTCTTCCATT	5016
DB	4388	CCGTGGATCTGCAGCGCTCGCATTAGTTTCTTCAATAAACCACTCTATCTTTCATCT	4447
QY	5017	TGATCACCATTGGCACACATGCTTCAGATACTACTTACAGTAGACA-----CAG	5064
DB	4448	TGATCACCATTGGGCGCACATGCTTCAGATACTCTCTTACACAGATACAGATCTGTCTCCAG	4507
QY	5065	GCTTACCCCTTGTCTCAGGTTCAAGAGACAGTGAAGGCTCATTTCCGATCTTCTTTCT	5124
DB	4508	GGCCGCGCTTGTCTGAGGGGTGAAGAGATAGTAGAGGAGGCTCGGTGTGCACTTGTCTTCT	4567
QY	5125	TTGCAGAAATTTCTCAATATACAGTGGCTCCATTGGGTGTGATATTCCTGGCTGGTATTT	5184
DB	4568	TTGTGGAAGTGTCCGACACACAGACGCGCTCTACTGGGTGGGTGCTCCCGGCTGTGTACC	4627
QY	5185	TGTGGGTCTCACTGAAGAAATGGCATCACCCCTTATCTTCGCTGTGTGCAATGTTTTTCC	5244
DB	4628	TGTGGCTCTCCCTGAGGAACGGCATCACCCCTTACCTCCGCTGTGTGCACTGTCTTTCC	4687
QY	5245	ACTATTTACTTTGGGGTAACTCCGCTCAGGAACTGTGCAATCAAAATCTGCAAGAGGAGAT	5304
DB	4688	ACTATTTACTTTGGAGTAGTCTCGGCTTGAAGAACTGTTTGGCAATTTCTGCTGAAAGGAGAT	4747
QY	5305	ACAGTGCACTGTGTAGCTACTTACTTTTACTCAAAATTTGTTCTCTCTCTTCCAGGAAT	5364
DB	4748	TCAGTGCACTGTGTAGCTACTTACTTTTACCACAAATTTGTTCTCTCTCTTCCAGGAAT	4807
QY	5365	ATTGGGATCTGTAAAGGCCCTTGTCCAGAGGTGGTGTGAGATCTCTGCCCTTACTAACT	5424
DB	4808	ATTGGGATACCAATGAAGCCCTTACTACAGAGGTGGTGTGGAGATCTCTGCCCTTACTCAAGT	4867
QY	5425	GTTTGAAGCAAAAAACACCGTGTGTCAAGGTACCTTAGAAAAAGAAATAGTTTTGNATAGAC	5484
DB	4868	CTTTGAAGCAGAAAATGCTGTGGTCAAGTACCCCTAGAAAAAGAAATAGTTTTGNATAGAC	4927
QY	5485	TTCTGTGATGATATAGCTGCTCTGTGATCAAGCTTCTCAATTTTCAGGTGCCACGGTCTG	5544
DB	4928	TTCTGTGAGGACTACAGCTGTCTTCTAAATCAGGCTTCTCACTTTTAGGTGTGCACGGTCTG	4987
QY	5545	CAGATGATGAGCGAAAGCATCTGTCTCTCTGCCCTTTTCTGTGGGGCTATACTATGTTCTC	5604

Db 1273 TATTGAGATGCCCTTGGCCGATTTTATGTCAGATGGTGTGTTTACAAGAGGCCAGATGGT 1332
Qy 1734 GAGAAATCCCTGTCTCATAGCAGGTTAATGCTTTGGGATGCAAGACTTTATTAAGGTGCC 1793
Db 1333 GAAAACTCTTCTAGTGGCAGAGCTGATGCTTAGTGATTCCAAATTTATGGAAGGTGCT 1392
Qy 1794 CGTAAGATCCCTTCATGAATTCATCTTCAGCAGTGTGTTTATGAGAGATGGAATACAAAAA 1853
Db 1393 AGGAGTGATATCATCAGTGTGTTATGAGCAGTCTGCTTATGGATTTGAAATACAGAAA 1452
Qy 1854 CTCTTTGCTATGGAATTTGTGAAGTATTATTAACAACTGCGAAGAAATATATCAGTGAT 1913
Db 1453 CTAATTTGCTGTGATTTGCAAAAAATTCACGACAGTTGCGAGAGATTTTATGGAGAT 1512
Qy 1914 GATCATGACAGAGTATCTCTATTAACCTGCACTTTTCAGTTTCAGATGTTTACTGTTCTTACT 1973
Db 1513 GATCAGGAGCAGAGTGTGCGTGACTCTCTATCTGTCCAGTTCCTCACCGCACCTACT 1572
Qy 1974 CTGGCTGCACATCTTATTTGAAGAGCAGAAATGTTATCTCTGTCAATTTACTTGAACCTCTGCTA 2033
Db 1573 CTGGCTGCATCTCATACAGAGAAATCTTGATGAGCATTTATCAATTAAGCTTTTATG 1632
Qy 2034 GAAGTTTT-ACCTGAGTACTTGGACAGGAACAATAAAATTCAACTTCCAGGGTTAT 2087
Db 1633 GATCAATTTGACACATCGAGATGCCAGGGCAGATTTCAAGTTTGAACGATACACTGCTTTA 1692
Qy 2088 AGCCAGGCAAAATTTGGGAAGATATATGCGATATATGATGACCTTAAGATATATCTGATC 2147
Db 1693 CAAGCCTTCAAAATTTAGGAGAGTACAGAGCCTTATTTTATGATCTCAAGATATGTTAAT 1752
Qy 2148 AGCAAAACCAATATGACAGAAAGATTAAGAAATGCAATGCTTCTTGAAGTTTTCGATCT 2207
Db 1753 AGCAAAACCAATGATGCTGAGTGTGAGCTGAGGCGAAGATTTCTTGAAGGGTTTGAATGCC 1812
Qy 2208 TTTTGTGAAGATTTCTTACCTGTATGAGGGAATGGAAGAAATCCGAAGACAGGTTGGGCAA 2267
Db 1813 TTTTGTGAATTTACTAAATATGATGACAGGATGGATCCAAATACAGTCAAGTAGGACAA 1872
Qy 2268 CACATGAGTGGATCTCTGATTTGGAGGCTGCAATGCTATACAGATGCAATTTGAAGAT 2327
Db 1873 CATATGAAATGGAACACAGATGGGAGCAGCCTTCACTACAAATGAAATTAACAT 1932
Qy 2328 ATTTTACTCATGTTCCAGAGTGGTGTGCTGTGATGAAGAACTCTTACTTGTGCTTAT 2387
Db 1933 GTCAATTCATATGATGACAGACTGGTGTGCTTCAGATGAAAGTGTTAATCGAGCTTAC 1992
Qy 2388 AAAGAAATGTCAAGCTGTGATGAGGTGCACTACAGTTTCATATCTAGTAGCAAGACA 2447
Db 1993 AAGAAATGCTCGCTGTACTGATGCAAGTGTCAATGCTGTATCTGATGGTGAACAGCA 2052
Qy 2448 GT---AGTACAAATCGTGGACATAGTTTGGAAACAAAGTCCCTACAGAGTATCTGAGAT 2504
Db 2053 ATCACATTAAGCATTTGTGGACATTCAGTGGAAACATATCAGATATCTGTGTTTCCCAAGAA 2112
Qy 2505 CTGTGAGCATACATCTGCACTCTTAGGACCTTGTGCTTCTCATGTAGTTTAAAGC 2564
Db 2113 AAAGTTAGCATTCACCTCCAGTTTCTCGCTTACTTGCAGGTTTACATGTATTTAAGC 2172
Qy 2565 AGGCTGGGTGTGTTTCAAGACTGATGAAATTTGTGTCTTTTGAAGCACTTCAAGTAGAG 2624
Db 2173 AAAGTGAAGTGGCATATAAATTTCCAGAGCTCCCTACCTTAAGTGAACCTTAGCCACCC 2232
Qy 2625 GTACTAGTGGAAATATCTCTTATAGTTGTCTGGTGTGTTGCTGCGCAGGTTGTTGCTGAGATG 2684
Db 2233 ATGTTGATAGAAACACCTCTTAGATGTCTTGTCTGTGCTGCGCAAGTACATGCGGGAATG 2292
Qy 2685 TGGCGAAGAAATGGAATCTCTTATATAGCAGGTTTATTAATCAAGATGTTAAGTGC 2744
Db 2293 TGGAGAAATATGGGTCTCTCTAGTAAACACAGATTTATTAATCTACAAATGTGAAATGC 2352
Qy 2745 AGAGAAAGAAATGATGATAAAGATATCATCATGCTTTCAGATTTGGTGTGATCTTTTATGGAT 2804
Db 2353 AGACGTGAGATGTTTGAACAAGATGTAGTAAATGCTTTCAGACAGGTTCTCCATGATGAT 2412

Qy 2805 CCCAATAAGTTTCTTGTACTGGTACTTTCAGAGGTATGAACTTTCGCGAGGCTTTTAAACAAG 2864
Db 2413 CCAAAATCATTTCTTGATGATCATGCTCAGCGCTTTGAACTTTATCAGATTTTTCAGTACT 2472
Qy 2865 ACCATATCTACAAA-----GACCAGGATTTTGAATTAACAATAT 2903
Db 2473 CCAGACTATGGAAAAAGATTTAGTTCTGAGATTAACCCATAAGGATGTTGTTTCAGACAAC 2532
Qy 2904 AATACATTAATAGAAAGAAATGCTTCAGGTCTCTCATCTATATTTGTTGGGTGAGCGTTATGTA 2963
Db 2533 AATACTCTTAATAGAAAGAAATGCTATACCTCATTAATAATGCTTGTGGAGAGAGATTTAGT 2592
Qy 2964 CTTGGAGTGGGAATGTGACCAAGAGAGTCAATAGAGAGTCAATGAGAGAAATCATTTCACTTGCTT 3023
Db 2993 CTTGGAGTGGCAGGTAATGCTACAGATGAATCAAGCGAGAGATTTATCCATCAGTTG 2652
Qy 3024 TGCAATGAACCCATGCCACACAGTCGCAATTCGCAAAATTTTACCTGAGAATGAATAAT 3083
Db 2653 AGTATCAAGCCTATGGCTCATAGTGAATTTGTAAGTCTTTTACCTGAAGATGAGAACAA 2712
Qy 3084 GAAATCGCTTATAGAAATGCTATPAACAAGATGGCCACATTTTAAGAAACACAGGTGTATCA 3143
Db 2713 GAGACTGGCATGGAGAGTGAATCGAAGCAGTTGCCCATTTTCAAGAAACCTTGGATTAACA 2772
Qy 3144 GGCCATCGAGTTTATGAATTAAGATGAATCACTGAAAGACTTTCATATGCTACTTTTAT 3203
Db 2773 GGACGAGCATGTATGAATCGAAACCAAGATGTGCCAAAGAGTTTCACTTGTATTTCTAT 2832
Qy 3204 CATTACTCCAAAACCCAGCATAGCAAGCTGAAACATATGCAAGAAAGAGAGAAAAACA 3263
Db 2833 CACTTTTCAAGGGCAGAACAGTCCCAAGCAGAGAGCGCAACCGAAAAATTTGAAAGACAA 2892
Qy 3264 GAAAAAAGATGAAGCAATTTGCCGCCACCACTCTGAAATTTCTGCTTCTGCTTTCAGC 3323
Db 2893 AATAGAGAGATACAGCACTCCACCTCCGCTGTGCTTCCATTTCTGCTTCTGTTGCA 2952
Qy 3324 AAAGTGATTAACCTTCTCACTGTGATATCATGATGTACATTTCTCAGACCGTATTTGAG 3383
Db 2953 AGCTGGTTAACTATTTTGGCTGAGATGTCTGTTGTGATCATGTTGGGAAACATTTCTGCA 3012
Qy 3384 CGGCAATATAGACACAGATTTCTAACTTGTGGACCGAAGGATGTCTTCAATGGCTTTTCT 3443
Db 3013 TGGCTGTGGAACATAATGGATATGCTGTGTCAGAGTCCATGCTGCAAAAGGTTTACAT 3072
Qy 3444 ATTCTGGCATTTGGGTTTCTTAGAAGAGAGCAAGCTTCAAAAGCTTCTGAAAG---AA 3500
Db 3073 TTAATTTGGCATGGCACTACAAAGAAAGAAAAACAATTTAGAGAAATGTCAAGAAAGCAT 3132
Qy 3501 GAAGTAAACATTTGACTTTTATCATAGGCTTCAAGATTTGGGAAGTTTCAAGCCATGAAT--- 3557
Db 3133 GTAGTAAACATTTACCTTCACTCAGAGATTAACAACTGTGTGAGCGCCAAAAAATTTCT 3192
Qy 3558 -----ATACAAATGCTTTTGGAAAAAACTCAAGGAAATTTCCAGTTTAGAAGCGCAGAAG 3611
Db 3193 CCTAGCATACTAGCTATGCTGGAACACTACAAATGCTCTCCCTACCTAGAGTCCACAAA 3252
Qy 3612 GACATGATAACGTGTGATCTTCAGATGTTTGAACAGTGAAGCGGATTAAGAGAAAAATCT 3671
Db 3253 GACATGATTCGTTGGATATTTGAAGACTTTTAATGCTGTATAAAAAAGATGAGGAGAGTTCA 3312
Qy 3672 TGTTTAAATTTGACCAACCATCAGGATCGGAATCTTATTAAGATGATGAGATTTACTCAT 3731
Db 3313 CCTACCAGTC-----CGTGGCAGAGACAGAGGAACCAATTAATGGAAGAGATTTCAAGG 3365
Qy 3732 GATAAGAAAAAGCAGAAACGAAAAAGAAAGCTTGAAGCTGTAGGCTACATCGCCAGAAG 3791
Db 3367 GACAAACAAAGCTGAGAGGAAGAAAAACAGAGATTTGCCAGACTGCGCAGAGAAAAAG 3426
Qy 3792 ATCATGGCTCAGATGCTGCTGCTTACAGAAAAAATTTTCAATTTGAACCTCAATAACTCATGTAT 3851
Db 3427 ATCATGGCTCAGATGCTCTGAAATCGAGCGCATTTTATTTGTAAGAAAAACAAAGAACTCTTT 3486

3852 GACATACATCAGAAATGCTGGGAAAGAGATTCATATATGGAGGAGAGGACCCCA 3911
3877 CAGCAGACATTAAGAACTGGATGCTCAACCTCTGCTGTTCTTGATCA-----TAGCCCT 3540
3912 GCAGTCAGTACTACTAGAAATGCTTTGGGTCTCTAAACGGGTCCATCTGTACTGAA 3971
3541 GTGGCTTCAGATATGACACTTACAGACTGGGTCCACACAACTCAGGTTCTGTGACAA 3600
3972 AAGGAGGTGCTGACGTGCATCTTTGCAAGAGAAACAGGAGGTGAATAATAGAAATAAT 4031
3601 AGACAAATTCGTTACATGATATATGTTCAAGAGGAGCAAGAGTTAAAGTGGAAAGCAGG 3660
4032 GCCATGATTAATCGGCTGTGTCCAGAAATCTACTGCTTAAACCCAGCAGAGGGGAAA 4091
3661 GCAATGGTCTTGGCAGCATTTGTTGAGAGATCACTGATTAATCAAAAACAGAAATGAA 3720
4092 CCCATAGAACTCTCAGGAGAGCCCTAGACCCACATTTTCATGATCCAGACTTGGCATAT 4151
3721 TTTATTCAGATCCAGAAA-----TATGATCATTTATTCATGACCCCTGATCTGTCTGT 3777
4152 GGAATTTATACAGGAGCTGTGTGTCATGTATGACGAGTGCTGTGCGAGAGTATTTT 4211
3778 GGAACACACACTAGTAGTGTGGGCACATATGATGATGCCATTTGTCGCAAGGTATTTT 3837
4212 GAAGCTGTACAGCTGAGCTCTCAGCAGCGCATTCATGTTGACCTTTT-----T 4259
3838 GATTCGGTTCAAGCTTAAGAAACAGCAGGAGCAAGATTAACGCTTACATACGAGCTAT 3897
4260 GACTTGGAAAGTGGAGATATCTTTGCCCTCTTTGCAATCTCTGCAATCTGATGATC 4319
3898 GATGTAGAAACGGAGAAATCTTTGCCCTTTGTGAAAGTGTGAGTAATCTGTTAT 3957
4320 CCCATTAATCTTCTTCAACCTCAAGAGATAACAGTGAAGATGACAGATGCTCTTGCTCAA 4379
3958 CC-----TCTGCTCTCTCCAGAAATATTTTAAACAGGTTAAATTTTTCAGAC 4011
4380 CTTTGTGACCTGGCAGGTGATACAGACTGTTCTGGCCAGAAATACAGGTTATATATA 4439
4012 CAACAAATCTGACTCAGTGGATTAAGAAACAATATCTCAGCAAAATAAAGCATTAACGTTT 4071
4440 AGACATGCTAAGAGGAAACCCATCTCTATTTCTTTTAAACAGGATGGAGATTC 4499
4072 CTTAGGAAGAGAAAGTATCTCTTAATATGCTCTACAAGAAATTCAGAAATGTGAT 4131
4500 ACTTTGAGTTCCTCATCTGAGTTTGGCGTTTGGAGTCTTCGATTAATAATATTCAAAT 4559
4132 GAAATACAGCTCCCTGAAGGTTCAAGGCTGATTTCTGCTCTAGATCCCTTATCTGAG 4191
4560 AGCATCAAGGAAATGGTTATCTCTTTGGCCACAAATTTATAGAAATTTGGAATGAAAGTG 4619
4192 AGCATAAAGAAATGCTAAACGACATTTGGAACTGCTACTACAGGTTGGGACTTAAAGGT 4251
4620 CCACCTGATGAAGGATCTCGATGCCATGCTGACCTGGAGCAGCTGGCTTCACT 4679
4252 CATCCAAATGAAGAGGATCTCGTGTCCCAATATGTTGGGGTAGCTGGCGGTACACC 4311
4680 ATCCAGGCAATGAATAATCTATTTGGAGATGAAGGAAACCTCTGTTTGGAGCACTTCAA 4739
4312 ATCCAAAGCATAGAAAGAAATTTGATGATGAGATTAACCAATGTTTGGTCTTACT 4371
4740 AATAGGACGATTAATGGTCTGAAGCATTAATGACAGTTTGGAGTTGCAAGAGGATACC 4799
4372 TGCAGCTGGATGACTGTCTTAGGTCAATTGACGAGATTTGGCGGACACACTGGACAGTG 4431
4800 TGTCTCAGTCTCTGATACAGAAACATCTGTTGCTCTTCTATCAGTTGTTCTTCTTAAC 4859
4432 GCATCAGTTTCAGTGGTGCAGGACATTTTGTAAACTTTTGTGATCACTGGTGCCTAAT 4491
4860 ATAAATCAGAAAGATACACCACTCTCTGCTATATAGATCTGTTTCAATGTTTGGTGGT 4919
4492 GACAGCCATGAGAACTTCATGATGATATAGATATGACATGTTTCAATTTATTTGGTGGC 4551
4920 GCTGTGTTAGCATTCCTTCTGTTGTTGGATGACCCCTGTTGATCTGACGCTTCTTCA 4979

4552 TTGGTCTTGCATTTCTGCGTTCAGTG-----TCAGATTTTTCAGGG 4596
4980 GTTAGTTCTTCTCTATAACACCTTTATCTCTTCAATTTGATCACCATGCGCACATGCTT 5039
4597 ATCAGCTTGGCACTGGAGACTTTCACATTTTCCATCTGGTTACTATATGGCACATCATA 4656
5040 CAGATATCTATCTTACAGTAGACACAGGCC-----TACCCCTTGTCTCAGGTTCAAGAAGACAGT 5096
4657 CAGATCTTACTTACCTCATGTACAGAGAGAAATGGCATGGATCAAGAAAAATCCCCCTTGT 4716
5097 GAAGAGGCTCATTCGGCATCTTCTTCTTTCGAGAAATTTCTCATATATACAAGTGGCTCC 5156
4717 GAAGAGAAATCAGCAGTTCTTGTCTTGTATAAAAACACTTCCACAGTATACGGGAAAGTGC 4776
5157 ATTGGGTGTGATATTCCTGCTGCTGATTTTGGGTCTCACTGAAGAAATGGCATCACCCCT 5216
4777 TTGAAGAAATACCATCCGCTGCACTCTGTGGAGGAGTGTGAGAGCTGGATCATGCT 4836
5217 TATCTGCTGTGTGCAATGTTTTTCCATTTATCTTGGGTTAACTCCGCTGAGAA 5276
4837 TTCTGAGAGTCTTCTGCTTTATTTTTCATTTAAATGGAGTTCTTCCACCACCGAC 4896
5277 CTGCATACCAATTTCTGAGAGGAGATACAGTGCATCTCTGATCATCTATCTTACT 5336
4897 ATTC-----AGTTCTTGGAAACAGCCATTTTGAAACATTTATGTAGCTATCTTCCCTACA 4953
5337 ACAATTTGTTCTGCTCTTCCAGAAATATTTGGGATCTGTAAGGCCCTTGTCTCAGAGG 5396
4954 AACAACTCATTTGCTTTTTCAGAAATATGATGATTAATGAATTTCACTGATGAAAGT 5013
5397 TGGTGTGAGATCTGCTCTTACTAACTGTTTGAAGCAAAAAACACCGTGGTCAAGTAC 5456
5014 TGGTGGCTAAACAGTGAAGTTAAAGATATCTAGAGGTGAAAGAGATGCTATAAGATAT 5073
5457 CTTAGAAAGAAATAGTTTGTATAGAGCTTCTGATGATCTGATCTGCTCTCTGATCAA 5516
5074 CCAAGAGATCTAAACAAATTAATAACCTTCCAGAGGATTAACAGCCTCATTAATCAA 5133
5517 GCTTCTCATTTTCAAGTCCACCGTCTGCAATGATGAGCGAAAGCATCTGCTCTCTGC 5576
5134 GCATCCAAATTTCTGTCGCGCAATCAGGTGGTGTATGAGAGCAGGCCCCAACTCTGTGC 5193
5577 TTTTCTGTGGGCTATATATGTTCTCAGAACTTTGCTGCCAGGAAATTTGTGAACGG 5636
5194 CTTGTGTGCGATCTCTGCTGTCTCCAGAGTTACTGCTGCCAGACTGAACCTGGAAGGG 5253
5637 GAAGAGTTGAGCTTGCATTTTTCACGCACTTCACTGTGAGCGGAGTCTGCAATTTTC 5696
5254 GAGGATGTAGAGCTTGCACAGCTCACACCTTCTCTGTGGCTCTGGAGTGGGCACTTTC 5313
5697 CTAAAAATCAGAGAAATGCGAGTGTCTGTTGAAAGTAAAGCCAGAGGCTGTGCTCTAT 5756
5314 CTGAGATACGGAAATGTCTGCTGCTATTTTGTGCAAAACCAAGGCTGTTTTTAT 5373
5757 CCAGCTCTTACTTTGGATGATATGAGAAACAGACCCCTGGCTGGAAGAGGGGCAACCCC 5816
5374 TCTCTCTTCTTACCTTGTGATGACTATGSGGAGACCCAGGAGCTCAGACGGGAAATCCT 5433
5817 CTTCAATTTCTCTGAGCGGTATCGGAGCTCCATTTGTTGTTGTTGTTGTTGTTGTTGTT 5876
5434 TTACATTTATGCAAGAGCGATTCAGAAAGATTCAGAAAGCTCTGGCACAACACAGATGTC 5493
5877 ATAGAAGAGATTTGCTAGGAGCAAGAGACTAATCAGATGTTTATTTGGATTCAACTGGCA 5935
5494 ACAGGAAATTTGACATGACAGAAACCAATCAGACACTGTTGGCATTTGACTGGCA 5552

RESULT 6

US-09-724-126A-5

; Sequence 5, Application US/09724126A

; Patent No. 6706505

; GENERAL INFORMATION:

; APPLICANT: Han, Hui-Quan
; APPLICANT: Kwak, Keith
; TITLE OF INVENTION: Human E3 Alpha Ubiquitin Ligase Family
; FILE REFERENCE: 01017/35966A
; CURRENT APPLICATION NUMBER: US/09/724.126A
; CURRENT FILING DATE: 2000-11-28
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 6089
; TYPE: DNA
; ORGANISM: Mouse
US-09-724-126A-5

Query Match 19.6%; Score 1233.6; DB 4; Length 6089;
Best Local Similarity 54.7%; Pred. No. 0;
Matches 2870; Conservative 0; Mismatches 2269; Indels 105; Gaps 17;

QY 769 TGGCATCTTGGTGGATCAGCAAGTTGATTTTATACCTGCTTCTTGATCATTTGGCAC 828
DB 836 TCGCAGGAGATGGCTGCAAGCAACCGACCTCAACAGAGAAAGTGTAACAGCATTTAGCCC 895
QY 829 AATTGGTCCAGAAATTTACTTTGCTGAAATGGACCCAGACTTGGAAAGCAGGAGAAA 888
DB 896 ACTGTGTGCCAAATCTACTTGCCTGGGGCCCTAACCCCTTCCCTCAGAAAGAGACAGC 955
QY 889 GTGTACAAATGTCAATATTTCACTCCACTGGAAATGGTACTTATTTGGAGAAGATCCAGATA 948
DB 956 TGGCAGCAGCATCTCTGCTGGACCGATGGAGTGGTACATCTGCGCTGAAGACCTGGCG 1015
QY 949 TTGCTGTAGAGAAATGAAGCAC---AGTGAGCATTTTACGCTTTGCTGGAGGGTTTCA 1005
DB 1016 TGGGATTTCCAAAGCTCGCAGGCGCAACAAGCCTTCTCACTCTGTGGCGAGTGTTA 1075
QY 1006 AAGTGGAGAGACACCTATTCTTCGAGGATTTGCAATTTGATCCACATGTGTACTCT 1065
DB 1076 AAGTGGGGGAACCTTACATCTCTGCGAGACTGTGCGAGTTGACCCACCTGTGTAT 1135
QY 1066 GTATGGAATCTTCCAGGACAGTGTTCATAAAATCATCGTTTACAGATGCATCTCTA 1125
DB 1136 GCATGGAGTGTCTCTGGAGATTCATAGAGACCATCGATATAGGATGACCATCGG 1195
QY 1126 CTGAGGAGGGTCTGTGCTGTGAGACACAGAGGCAATGAAACTGGCCCTTTTGTG 1185
DB 1196 GAGGAGGGGGTCTGTGCTGTGACACTGAGGCGTGGAAAGAGGAGCCTTTACTGCC 1255
QY 1186 TAAATCATGAACCTCGAAG---AGCAGGTACTATAAAGAGAAATTCACGCTGTCCGTGA 1242
DB 1256 AGAAGCAAGCTCAGCAGCTCTGAAGTTGTGGAGGAGAGATCCTCTTGTGCATCTAT 1315
QY 1243 ATGAAGAGTAATTTGCAAGCCAGGAAATATTTCTTCAGTGATGATAAAATATGCTGAG 1302
DB 1316 CAGAAGATGTATGCCAGAACTTACACATTTTGTCTATTATGTTTCGATATGCAGTAG 1375
QY 1303 AAATGACTATATGGGAAGAGGAAAGAAAGCTGCTCTGAACTCCAGATAGAGGAGAAA 1362
DB 1376 ATATACTGACCTGGGAAAGAAAGTGAATTTGCTGGAAGACTTAGAAGTGGCAGAGA 1435
QY 1363 ATGAAGATATCTATTGTGCTCTTTTCAATGATGAACACCATTTCAATGACACAGCTAT 1422
DB 1436 GTGACACCTTACTCTGCTATGCTGTTTATGATGAGGTTTCAACCTTATGAGCAAGTCA 1495
QY 1423 ACAGCCTTACAAGAGCTCTTGACTGTGAGCTGCGAGAGGCCAGTTTGCATACCACTGCCA 1482
DB 1496 ATACCTTTCAGAAAGCTGTGAACCTGTACACAGAGAGAAAGCAATTTGGCTTTGCAAC 1555
QY 1483 TTGACAAAGAGGCTGTGGGCTGTGTTAAAGCGGAGCTTATGCTGCTTCCAGGAGCAA 1542
DB 1556 TTGATCGAGATGGCCGTGAGGCTGTCCGATATGGAGATTTCCAGTACTGTGATCAAGCA 1615
QY 1543 AGGAAGATATAAGAGTCAATTCAGAAATGTCTCTCAACATCCACTTTCATGTAGAAGTAT 1602

DB 1616 AGACAGTCAATTTGTGAGGAACACACAGCAGACAGACCAAG---CCGCTCAAAGTTTCAAGTTA 1672
QY 1603 TACACTCAGAGATATATGGCTCATCAGAAATTTGCTTTGGCTCTTGGTCTCTGGATGAACA 1662
DB 1673 TGCACCTCTCCGTGGCTGCTCATCAGAAATTTGGTTTGAAGCTCTCTCGTGGCTGGAA 1732
QY 1663 AAATTATGAGCTATTCAAGTGACTTTTAGGCGAGATCTTTTGCAGCATGCTTATAGAGAAG 1722
DB 1733 GTGTTATGGATCTCAGATGGCTTCCGAGGATTTTGTCTCAAGTTGGATTACAGAGAG 1792
QY 1723 AACCTGACTCGGAGAAATCCCTGTCTCATTAAGCAGGTTTAATGCTTTTGGGATGCAAGCTTT 1782
DB 1793 GTCCAGATGGCAAACTCTTCTCTGCTCGACAGACTGATGCTTAATGATTTCCAAATTA 1852
QY 1783 ATAAAGTGGCCGTAAGATCCCTCATGAATTTGATCTTTCAGCAGTTTATTTATGAGATGG 1842
DB 1853 GGAAGGGGCTAGGAGTGTGTATCACCAATTTTTCATGAGCAGCTGCTCATGAGCCTCA 1912
QY 1843 AATACAAAAAATCTTTTGTGTAAGTATTAATAACAACTGACAGAAAGAAAT 1902
DB 1913 AGTATAAGAGCTGTTTCGCTTCCGATTTGCTAATAAATTTACCGCAGTTTGCAGAGGAT 1972
QY 1903 ATATCAGTGTATCATGACAGAAAGTATCTCTATAACTGCACTTTTCAGTTTCAGATGTTA 1962
DB 1973 TTATGGAGGATGATCAGCAGCGGCGAGTGTGGTGAAGTCTGCTCTGTCTGTCCAGTTCTCA 2032
QY 1963 CTGTTCTCTCTGGCTCGACATCTTATTGAGAGCAGAAATGTTATCTCTGTCTATTACTG 2022
DB 2033 CCGCAGCAGCTGTGGCGGAAATGCTCTCTCAGAAAGAGAACCTGATGACCGTTATCATTA 2092
QY 2023 AAATCTGTCTAGAAAGTTT-----ACCTGAGTACTTTGGACAGGAAACAATAAATTTCAACT 2076
DB 2093 AGGCTTTTCATGAGCCATTTGAAACACAGAGATGCCAGGCGAGATTCAGTTTGAACGCT 2152
QY 2077 TCCAGGGTTATPAGCCAGGACAAATTTGGGAAGATATATGCAAGTATATGTGACCTAAAGT 2136
DB 2153 ACATGCGCTCCAGCCCTCAAGTTTCAGGAGATCCAGAGCTCATCTTAGATCTCAAGT 2212
QY 2137 ATATCTGTATCAGAAACCCACAAATATGACAGAAAGATTAAGAAATGAGTTTCTTGAAG 2196
DB 2213 ATGTATTTGATTAGCAAAACCAACGAGTGGTGCAGATGAGCTGAGGCGAGAAATTTTACAA 2272
QY 2197 GTTTTCGATCTTTTTCAGAAATTTCTTACCTGTATGAGGGAATGGAAGAAATCCGAAGAC 2256
DB 2273 GGTTCGATGCTCTTGTGAATTTACTGAGTGCATGAGGGAATGGAACCGATCAGCGTC 2332
QY 2257 AGTTTGGGCAACACATTTGAAGTGGATCTGATTTGGGAGGCTGCCATTTGCTTATACAGATGC 2316
DB 2333 AGTGGGACAGCACATTTGAGATGGAGCCAGAGTGGGAAGCAGCTTTCACACTGCAGATGA 2392
QY 2317 AATTGAAGATATTTTACTCATGTTCCAGAGTGGTGTGCTTGTGATGAGAACTCTTTAC 2376
DB 2393 AGCTGACACAGCTCATCTCAATGGTGCAGGACTGGTGTGCTCTGGAAGAAAGTGTAA 2452
QY 2377 TTGTTGGCTTATAAGAAATGTCACAAAGCTGTGATGAGTGCAGTACCAAGTTTTCATCTA 2436
DB 2453 TTGAAGCTTACAGAAATGCTGCTGTGCTGTCACAGTGTCTATGGCGGATTTACTGATG 2512
QY 2437 GTAGCAAGACAGT---AGTACAATCGTGTGGACATAGTTTGGAAACAAAGTCTTACAGAG 2493
DB 2513 GTGAACAGCCAAATCACACTCAGTATTTGTGACACACTCGGTGGAAACCAATCAGATCTGTG 2572
QY 2494 TATCTGAGGATCTTGTAAAGCATATCTGCGCACTCTCTAGACCCCTTGTGCTTCTTCAAG 2553
DB 2573 TTTCCCAAGAAAAAGTTAGCATTCACCTCCCAATTTCTCGCTTGTGTCAGGTTTGCATG 2632
QY 2554 TACGTTTAAAGCAGCTGGTGTCTGTTTCAAGACTGCAATGAAATTTGTGCTTTTTCAGGACT 2613
DB 2633 TATTGTTAAGCAAAAGTGAAGTGCATATAAATTTCCAGAGCTCTTACCTCTTAAGTGAAC 2692
QY 2614 TTCAAGTAGAGGTAATAAGTGAATATCTTTTACGTTGTCTGGTGTGTTGTTGCCCAAGTTG 2673

Db 2693 TGAGCCCAACCATGTTGTAGAACATCTCTTAGATGCTTGTCTTATGTGCTCAAGTGC 2752
Qy 2674 TTGCTGAGATGGCGAAGAAATGGAATGCTCTCTTATAGCAGGTTTATTAACCAAG 2733
Db 2753 ATGCTGGATGTGGAGAGAAATGGCTCTCTAGTAAATCAGATCTATTAACCATTA 2812
Qy 2734 ATGTTAAGTCAGAGAGAAATGTATGAAGATATCATCATCTTCAGATTTGGTCAT 2793
Db 2813 ATGTGAATTCAGGCGAGAGATGTTGCAAGGACATAGTGTCTTCAGACAGGTGCT 2872
Qy 2794 CTTTAAATGATCCCAATAGTTCTTGTACTTCTGTAATCTCAGAGGTATGAATTCGCGAGG 2853
Db 2873 CCAATGATGACCCCAACCACTCTCTGATGATCATGCTCAGCGCTTGTGAATCTATCAGC 2932
Qy 2854 CTTTAAACAGAACCATATCTCAAAAG-----ACCAAGATTTGA 2892
Db 2933 TCTTCAGCACGCTGACTATGGAAGAGATTCAGTTCTGAGGTTACCCATAGGACGTCG 2992
Qy 2893 TTAACAATATATACATTAATAGAGAAATGCTTCAGGTCCTCATCTATATTTGGGTG 2952
Db 2993 TTCAGCAAGAACCACTCTGATCGAAGAGATGCTCTACCTCATCATGCTTGTGGGAG 3052
Qy 2953 AGCGTTATGTACTCGAGTGGGAAATGTGACCAAGAGAGGTCACAATGAGAGAAATCA 3012
Db 3053 AAAGATTCACCTGGGTTGACAGGTGGCTGCCACAGTGAATCAAGAGGGAGATTA 3112
Qy 3013 TTCACTTGTGTTGCAATGCAACCATGCAACAGTGCCATTGCCAAAATTTACCTGAGA 3072
Db 3113 TCCATCAGTTGAGCATCAAGCCTATGGCTCACAGTGAGTGTGTAAGTCTCTGCCTGAAG 3172
Qy 3073 ATGAAATTAATGAATCTGGCTTAGAGATGTGATTAACAAGTGCCGACATTTAAGAAAC 3132
Db 3173 ATGAAACAAGAGAGACCGGATGAGAGGCTCATCGAGTCCGTTGCAATTTCAAGAAAC 3232
Qy 3133 CAGGTGTATCAGGCCATGAGTTTATGAATGAATGAATGAATGAATGAATGAATGAATGA 3192
Db 3233 CTGGCTCAGCGGCGAGGATGATAGCTGAGCGAGGTGCCAAGAGTTCAACC 3292
Qy 3193 TGTACTTTTATCATTTCTCAAAACCCAGCATAGCAAGGCTGAAATATATGCAAGAGAAAC 3252
Db 3293 TGTATTTTATCATTTCTCCAGGCGAGCAGTCCCAAGGAGAGAGTCCAGCGGAAAT 3352
Qy 3253 GGAGAAACAGAGAAACAAAGATGAAGATTCGCGCCACACCACTCTGTAATTTCTGCC 3312
Db 3353 TGAAGAGAGAAATTAAGAGATACAGACATCCCTCTCCGCTTTGGCCACCGCTTCTGCC 3412
Qy 3313 CTGCTTTCAGCAAGTGATTAACCTTCTCACTGATATATGATGATGATGATGATGATGATG 3372
Db 3413 CTTTGTTCGCGAGTCTGGTTAAACATCTTGCAGTGTGACGTATGCTGTACATCATGGGAA 3472
Qy 3373 CCGTATTTGAGCGGCAATAGACACAGATTTCTAACTTGTGACCGAAGGAGTCTCCAAA 3432
Db 3473 CGATCTCGATGGCTGTAGACATCAGGGTCTGCTGTGTCAGAGTCCATGCTACAGA 3532
Qy 3433 TGGCTTTTCAATTCGCGATTTGGTTTATAGAGAGAGCAACAGCTTCAAAAGCTC 3492
Db 3533 GGGTCTGCTATTTGATCGGGATGGCTCCAGAGAGAGAGACCACTTGGAGAACGCG 3592
Qy 3493 CTGAGAGAGAGT---AACATTTGATTTTATCATTAAGCTTCAAGTTGGAAGTTTCAG 3549
Db 3593 TGGAAAGGCGACGTGACAGACCTTCACTTTCACACAGAGATTTCAAGCCTGTGTATGCAC 3652
Qy 3550 -----CATGAATATACAAATGCTTTTGGAAAACTCAAGGAATTCCTCCAGTTAG 3600
Db 3653 CACATACTCCCGAGATCTTAGCTATGCTGGAGACCTTGCAGAACGCCCTCCCTGG 3712
Qy 3601 AAGCCAGAGACATGATTAAGTGGATATCTTCAGATGTTTGAACAGTGAAGCGATTA 3660
Db 3713 AAGCCACAGACATGATGATGCTGGTGTGCTAAAGATGTTTAAATGAAGAGATA 3772
Qy 3661 GAGAAATCTTGTATTTAGCAACCATCAGGATCGGAATCTATTAAGATGATG 3720
Db 3773 GAGAGTGTTCATCCAGACGCTGTGGCC-----GAGCGGAGGGAACCATTAATGGAGG 3826

Qy 3721 AGATTATCTATGATTAAGAAAGAGAGAAACGAAAGAAAGAAAGCTGCTAGGCTAC 3780
Db 3827 AGAGCTCAAGAGACAAGGACAAAGCAGAGAGAGAAAGAAAGCCAGATCGCAGACTGC 3886
Qy 3781 ATGCCCAAGATCATGCTCAGATGCTGCTTACAGAAACATTTCAATTTGAAATCTATA 3840
Db 3887 GCCGGAGAGATCATGCCCCAGATGCTGAGATGCGGGCACTTCATTGACGAAACA 3946
Qy 3841 AACTCATGTATGACAAATACATCAGAAATGCTCGGAAAGAAATTCCTATTATGAGGAAG 3900
Db 3947 AAGAGCTCTCCAGCAGACCTTAGAGCTGGACACCTCTGCTGCTGCTCTTGAC---- 4002
Qy 3901 AGAGCAACCCAGCAGTCACTACTCTAGATTTGCTTGGGTCTTAAACGGGTTCAT 3960
Db 4003 --AGCAGCCTCCCGTTTCAGACGCGCTCTTACAGCACTGCGGCCAGCAGACACAG 4060
Qy 3961 CTGTACTGAAAGAGGAGTGTGAGCTGCTCATCTTTCGCAAGAGAAACAGGAGGTGAAA 4020
Db 4061 TCCCTGAACCCAGACAGTGTGTCACTGTATATATGTCAAGAGAGCAAGAGGTGACTG 4120
Qy 4021 TAGAAATATATGCTATGTTATCGGCTGTGCTCAGAAATCTACTGCTTAAACCCAGC 4080
Db 4121 TGGAGAGCAGGCGATGCTTGGCAGGTTTGTTCAGAGGTCAACGTTCTGTCAAAG 4180
Qy 4081 ACAGGGGAAACCCATAGAACTCTCAGAGAGGCCCTAGACCCACTTTTCATGAGTCCAG 4140
Db 4181 ACAGGACGAAACCATCGCGACCCAGAAAAATAT---GATCCATTAATTCATGCAACCCG 4237
Qy 4141 ACTTGGCATATGGAATTTATACAGGAGTGTGCTCATGTAATGCAAGCAGTGTCTGGC 4200
Db 4238 ATCTGTCTGTGGGACACACACTGCGAGCTGTGGCAGCTTATGCAATGCTTGTGGC 4297
Qy 4201 AGAAGTATTTGAAGCTGTAC-----AGCTGAGCTCTCAGCAGCGCATTCATGTTG 4251
Db 4298 AAGGTATTTGATTCGTTCAAGCAAGAGAGCAGAGGAGCAGCGGCTGCGCTGC 4357
Qy 4252 ACCTT---TTTGACTTGGAAAGTGGAGATATCTTTCCTCTTTTGGCAATCTCTGTGCA 4308
Db 4358 ACATAGCTACGATGTAGAGATGCGAGTTCCTCTGCGCGCTCTGTGAGTGCCTGAGCA 4417
Qy 4309 ATACTGTGATCCCATTTATCTTTTGCNAACCTCAAGATTAACAGTGAATGAGATG 4368
Db 4418 ACACGGTATCCCT---CTGCTGCTTCTCCAGGAGCATCTCAGCAGGAGGTTAA 4471
Qy 4369 CTCTTGTCTCAACTTTTGACCTGCGCAGCTGATACAGACTGTTCTGCGCAGAAATATCAG 4428
Db 4472 ATTTTCAGCAACACAGATCTGCAAGTGGACGAGGAGTAAACAGCAGATTAAGG 4531
Qy 4429 GTTATATATAGACATGCTTAAAGGAGAAACCCCAATTCCTATTCTTTTAAATCAAGGAA 4488
Db 4532 TGGTCCAGATGCTGAGGAGAAAGCAATGCTGCTGACACGCTCTCTTCAGAGGACACAG 4591
Qy 4489 TGGAGATTTCTACTTTTGGAGTTCCATTCATCTGAGTTTGGGTGAGTCTTGCATTA 4548
Db 4592 AAGCCATGAATATATACCTATCCCGAAGGCTTCAGGCTGATTTTATCTTAGGAAAC 4651
Qy 4549 AATATTCAAATAGCATCAAGGAATGTTTCTCTTTGGCCACAACAAATTTATAGAAATG 4608
Db 4652 CATATCTGATAGATTAAGAAATGTTAAACATTTTGGAAACGCTCTTACAGGTGG 4711
Qy 4609 GATTGAAAGTGCACCTGATGAAAGGATCTCCAGTCCCCATGCTGACCTGGAGCACT 4668
Db 4712 GACTGAAGTTTCACTCTAAATGAAGTGAACCCCGTGTGCCCATCTCTGTCTGGGGACCT 4771
Qy 4669 GCGCTTTCTATCTCAGGCAATGGAATCTATTTGGGAGATGAAGGAAAACTCTGTTTGG 4728
Db 4772 GTGCATACCATCCAGAGCATAGAAAGAAATTTTGAAGTGAAGAGAGAGGCTGTTTTG 4831
Qy 4729 GAGCATTCAAAATAGCAGCATATGCTGTGAAGCAATTAATGACGTTTCAGCTTGCAC 4788
Db 4832 GACCTCTGCTTGTAGACTGGAACGACTGTCTCAGGTGCTTAAACAGCGTTTGACAGCAGC 4891

Db 241 CCACACAGTGCCTTCCCAAAATTTACCTGAGATGAAAATTAATGAAACTGGCTTAGAG 300
Qy 3099 AATGTCTATAAACAAGTGGCCACATTTAAGAAACACAGGTGTATCAGGCCATGGAGTTTAT 3158
Db 301 AATGTCTATAAACAAGTGGCCACATTTAAGAAACACAGGTGTATCAGGCCATGGAGTTTAT 360
Qy 3159 GAATCTAAAGATGATCACTGAAGACCTTCAATATGTACTTTTATCATCTACTCTCCAAACC 3218
Db 361 GAATCTAAAGATGATCACTGAAGACCTTCAATATGTACTTTTATCATCTACTCTCCAAACC 420
Qy 3219 CAGCATAGCAGGCTGAACATATGAGAGAAAGAGAGAAACAGAGAAACAGAGATGAA 3278
Db 421 CAGCATAGCAGGCTGAACATATGAGAGAAAGAGAGAAACAGAGAAACAGAGATGAA 480
Qy 3279 GCATTCGCGCCACACACCTCTCTGAATTCGCGCTCTTTCAGCAAAAGTGAATTAACCTT 3338
Db 481 GCATTCGCGCCACACACCTCTCTGAATTCGCGCTCTTTCAGCAAAAGTGAATTAACCTT 540
Qy 3339 CTCAACTGTGATATCATGATGTACATTTCTCAGACCGTATTTGAGCGGCAATAGACACA 3398
Db 541 CTCAACTGTGATATCATGATGTACATTTCTCAGACCGTATTTGAGCGGCAATAGACACA 600
Qy 3399 GATTCCTAACTGTGACCGAGGGATGCTCCAAATGGCTTTTTCATATCTTGGCATTTGGT 3458
Db 601 GATTCCTAACTGTGACCGAGGGATGCTCCAAATGGCTTTTTCATATCTTGGCATTTGGT 660
Qy 3459 TTACTAGAGAGAGACAGCTTCCAAAGAGCTCTCTGAAGAGAGAGTAAACATTTGACCTT 3518
Db 661 TTACTAGAGAGAGACAGCTTCCAAAGAGCTCTCTGAAGAGAGAGTAAACATTTGACCTT 720
Qy 3519 TATCTAAGGCTTCAAGATTTGGAGTTGCGACCATGATATCAAAATGCTTTTGGAAAAA 3578
Db 721 TATCTAAGGCTTCAAGATTTGGAGTTGCGACCATGATATCAAAATGCTTTTGGAAAAA 780
Qy 3579 CTCAAGGATTTCCCGAGTTAGAGCCAGAGGACATGATTAACCTGATCTTCCAGATG 3638
Db 781 CTCAAGGATTTCCCGAGTTAGAGCCAGAGGACATGATTAACCTGATCTTCCAGATG 840
Qy 3639 TTGTGACAGTGAAGCGGATTAAGAGAAATCTTGTGTTAATTTAGTGAACACATCAGGA 3698
Db 841 TTGTGACAGTGAAGCGGATTAAGAGAAATCTTGTGTTAATTTAGTGAACACATCAGGA 900
Qy 3699 TCGGAATCTTAAAGATGATGAGATTTACTCATGATAAAGAAAAAGCAGAAAAAGA 3758
Db 901 TCGGAATCTTAAAGATGATGAGATTTACTCATGATAAAGAAAAAGCAGAAAAAGA 960
Qy 3759 AAAGCTGAAGCTGCTAGGCTAGATCGCCAGAGATCATGGC 3799
Db 961 AAAGCTGAAGCTGCTAGGCTTATCGCCAGAGATCATGGC 1001

RESULT 8
US-09-228-317-2
; Sequence 2, Application US/09228317
; Patent No. 6159732
; GENERAL INFORMATION:
; APPLICANT: Varshavsky, Alexander
; APPLICANT: Kwon, Yong Tae
; TITLE OF INVENTION: NUCLEIC ACID ENCODING MAMMALIAN UBRI
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kevin M. Farrell, P.C.
; STREET: P.O. Box 999
; CITY: York Harbor
; STATE: ME
; COUNTRY: US
; ZIP: 03911
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/228,317
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Farrell, Kevin M.
; REGISTRATION NUMBER: 35,505
; REFERENCE/DOCKET NUMBER: CIT-2001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (207) 363-0558
; TELEFAX: (207) 363-0528
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1001 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..999
; US-09-228-317-2

Query Match 15.8%; Score 996.2; DB 3; Length 1001;
Best Local Similarity 99.7%; Pred. No. 5.2e-281;
Matches 998; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2799 ATGATCCCAATAAGTCTCTGTACTTGGTACTTTCAGAGGTATGAACCTTGGCCAGGCTTTT 2858
Db 1 ATGATCCCAACAAGTCTCTGTACTTGGTACTTTCAGAGGTATGAACCTTGGCCAGGCTTTT 60

Qy 2859 AACAGACCATATCTACAAAAAGACAGGATTTGATTAACAATATACTAATAGAA 2918
Db 61 AACAGACCATATCTACAAAAAGACAGGATTTGATTAACAATATACTAATAGAA 120

Qy 2919 GAATGCTTTCAGGTCCTCATCTATTTTGGGTGAGCGTTATGTACCTGGAGTGGGAAT 2978
Db 121 GAATGCTTTCAGGTCCTCATCTATTTTGGGTGAGCGTTATGTACCTGGAGTGGGAAT 180

Qy 2979 GTGACCAAGAGAGGTCAATGAGAGAAATCATTTCACTTGTCTTGCATTTGAACCCATG 3038
Db 181 GTGACCAAGAGAGGTCAATGAGAGAAATCATTTCACTTGTCTTGCATTTGAACCCATG 240

Qy 3039 CCACACAGTGCCTTGGCAAAATTTTACCTGAGAAATGAAAAATATGAAACTGGCTTAGAG 3098
Db 241 CCACACAGTGCCTTGGCAAAATTTTACCTGAGAAATGAAAAATATGAAACTGGCTTAGAG 300

Qy 3099 AATGCTATAAACAAGTGGCCACATTTTAAGAGAACCGGTGTATCAGGCCATGGAGTTTAT 3158
Db 301 AATGCTATAAACAAGTGGCCACATTTTAAGAGAACCGGTGTATCAGGCCATGGAGTTTAT 360

Qy 3159 GAATCTAAAGATGATCACTGAAGACCTTCAATATGTACTTTTATCATCTACTCTCCAAACC 3218
Db 361 GAATCTAAAGATGATCACTGAAGACCTTCAATATGTACTTTTATCATCTACTCTCCAAACC 420

Qy 3219 CAGCATAGCAGGCTGAACATATCAGAGAGAAAGAGAGAAACAGAGAAACAGAGATGAA 3278
Db 421 CAGCATAGCAGGCTGAACATATCAGAGAGAAAGAGAGAAACAGAGAAACAGAGATGAA 480

Qy 3279 GCATTCGCGCCACACACCTCTCTGAATTCGCGCTCTTTCAGCAAAAGTGAATTAACCTT 3338
Db 481 GCATTCGCGCCACACACCTCTCTGAATTCGCGCTCTTTCAGCAAAAGTGAATTAACCTT 540

Qy 3339 CTCAACTGTGATATCATGATGTACATTTCTCAGGACCGTATTTGAGCGGCAATAGACACA 3398
Db 541 CTCAACTGTGATATCATGATGTACATTTCTCAGGACCGTATTTGAGCGGCAATAGACACA 600

Qy 3399 GATTCCTAACTGTGACCGAGGGATGCTCCAAATGGCTTTTTCATATCTTGGCATTTGGT 3458
Db 601 GATTCCTAACTGTGACCGAGGGATGCTCCAAATGGCTTTTTCATATCTTGGCATTTGGT 660

Qy 3459 TTACTAGAGAGAGACAGCTTCCAAAGAGCTCTCTGAAGAGAGTAAACATTTGACCTT 3518
Db 661 TTACTAGAGAGAGACAGCTTCCAAAGAGCTCTCTGAAGAGAGTAAACATTTGACCTT 720

Db 661 TTACTAGAAGAGAACACAGCTTCAMAAAGCTCTGAAGAAGAAAGTAAACATTTGACTTT 720
Qy 3519 TATCATAGGCTTCAAGATTGGGAGTTCCAGCCATGATATACAAATGCTTTTCGAAAAA 3578
Db 721 TATCATAGGCTTCAAGATTGGGAGTTCCAGCCATGATATACAAATGCTTTTCGAAAAA 780
Qy 3579 CTAAAGGAATTTCCCACTAGAGAGCCAGAGACATGATAAGCTGTGATATCTTCAGATG 3638
Db 781 CTCAGGAATTTCCCACTAGAGAGCCAGAGACATGATAAGCTGTGATATCTTCAGATG 840
Qy 3639 TTTGACACAGTGAAGCGAATTAAGAGAAAAATCTTGTATTAATGTAGCAACACATCAGGA 3698
Db 841 TTTGACACAGTGAAGCGAATTAAGAGAAAAATCTTGTATTAATGTAGCAACACATCAGGA 900
Qy 3699 TCGGATCTATTAAGAAATGATGATTAATCTATCATGATAAGAAAAAGCAGACGAAAAAGA 3758
Db 901 TCGGAATCTATTAAGAAATGATGATTAATCTATCATGATAAGAAAAAGCAGACGAAAAAGA 960
Qy 3759 AAAGCTGAAGCTGCTAGGCTACATCGCCAGAGATCATGGC 3799
Db 961 AAAGCTGAAGCTGCTAGGCTTCATCGCCAGAGATCATGGC 1001

RESULT 9

US-09-270-767-14761
; Sequence 14761, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 14761
; LENGTH: 670
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-14761

Query Match 1.8%; Score 113.4; DB 4; Length 670;
Best Local Similarity 53.4%; Pred. No. 3e-22;
Matches 287; Conservative 0; Mismatches 241; Indels 9; Gaps 2;
Qy 2943 ATTGTGGTGAGCGTTATGTACTTGGAGTGGGAAATGTGACCAAGAGAGAGGTCAATG 3002
Db 2 ATCATTTGAGAACGTTGGATGCTGGCGTTTCGATGGTGACCGAAGAGGATCGCTGCGC 61
Qy 3003 AGAGAAATCATTCACCTTGTGCAATGCAACCCATGCCACACAGTGCATTCGCCAAAAT 3062
Db 62 AAGGAGATCATCCAGCTGTGTATCAAGCCATCTCACACTCGAGTTAGTGGCGCC 121
Qy 3063 TTACCTG-----AGAAATGAAATTAATGAAATGCTTTAGAGAATGTCTATAACAAAGTG 3116
Db 122 TTGCCAGATGCAATAGCGGAAATAGCGAACAGTTTTTCGAGGAGTTATCAACAGGTT 181
Qy 3117 GCCAATTTAGAAAAACGAGGTGATACAG---GCCATGAGGTTTATGAATCTAAAGATGAA 3173
Db 182 GCGGTGTTCAAGAAACCGGTTGGAGCCGACGACGAGGGGGTGATCAAACTAAAGGAACAT 241
Qy 3174 TCATGAAGACTTCAATATGTACTTTTATCAATTAATCTCCAAAACCCAGCATAGCAAGGCT 3233
Db 242 TTGCTTAAGGAGTTCAACATGTACTTCTATCACTACACAAAGAGGACAAGTCCAGGCC 301
Qy 3234 GAACATATGACAGAAAAAGGAGAAAAACAAGAAAAACAAGATGAAGCATTTGCCGCCACCA 3293
Db 302 GAGGAATCTGAACGGGAGCGCCGAAGGCCAAGAGAGAGCTTGTCTGTGCGCGCCACCA 361
Qy 3294 CCACCTCTGATATTCGCGCTGCTTTTCAGCAAAAGTATTAACCTTCTCAACTGTGATATC 3353
Db 362 ATGTTGCCAAAACCTAACACCGCTTCAACCCATGGCCCAACATCTCGAGTGGCCCTGTG 421

Qy 3354 ATGATGTATCATTTCTCAGACCGCTATTTAGCGGCAATAGACACAGATTTCTAATCTGTG 3413
Db 422 TTCTCAACATCTGCAGCTTGAATATGAAAAGGCTTTTAAATGCTTACAGTCGAGCTTT 481
Qy 3414 ACCGAGGAGTGTCCAAATGGCTTTTCATATTTCTGCAATTTGGTTTACTAGAAGAG 3470
Db 482 ACCGAATCGCATCTTCAAAAAGGCTCTTCACTGCTGGGTTATGCCATTCAGGAGGAG 538

RESULT 10

US-09-513-999C-936/c
; Sequence 936, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961

; FILE REFERENCE: 59, US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 936
; LENGTH: 366
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 37..366
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 8
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 9
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 317
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 318
; OTHER INFORMATION: k=g or t
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 94
; OTHER INFORMATION: Xaa=His or Leu or Pro or Gln or Arg
US-09-513-999C-936

Query Match 1.7%; Score 105.6; DB 4; Length 366;
Best Local Similarity 82.7%; Pred. No. 3.9e-20;
Matches 134; Conservative 0; Mismatches 24; Indels 4; Gaps 1;
Qy 123 AAAGCTTATCATCTGTCTTCCACTTATPCCAACAAAGCTGTATGGCCACTGCGCTGTGCGG 182
Db 173 ACAGGCTTATCATCTATCTTCCACTTATCCAAAGCGCTACGCGCACTTCTCTGTCCCG 114
Qy 183 CACTTGAACACCGCCAGCCCACTACTGCTCCACTACCACTGCTTCTCCACCCTGA 242
Db 113 CAGTTGAAGACCAACAGCCCACTACAGCTCCGTTACCAACCGCTTCTCCCGCCTGT 54
Qy 243 TCAGTGTCTGCTGCCATCTTATCGCTTCTGCTGTTC 284
Db 53 ACAGC-----TGCCGCGCCATCTTGTCCGCTCTGCGCTCTCC 16

RESULT 11

```

; LOCATION: 177..605
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 8
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 9
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 49
; OTHER INFORMATION: s=g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 199
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 8
; OTHER INFORMATION: Xaa=Ala or Asp or Gly or Val
US-09-513-999C-937

Query Match          1.6%; Score 103.2; DB 4; Length 605;
Best Local Similarity 81.5%; Pred. No. 2.8e-19;
Matches 132; Conservative 1; Mismatches 24; Indels 5; Gaps 1

QY 123 AAAGGCTTATCATCTGTCTTCCACTTATCCAAAGCTGTATGGCCACTGCGCTGTGCGG 182
Db 172 ACAGGCTTATCATCTATCTTCCACTTATCAAGCGCGGTACGGCCACTTCTCTGTGCGG 113
QY 183 CACCTGGAACCAAGCCAGCGCCCACTACTGCTCCACTACCACTGGTGTCTCCACCCCTGA 242
Db 112 CAGTTGGAAGCAACCAACCAAGCGCGGTACAGCTCCGCTACCAACGCTTCTCCGCGCTGT 53
QY 243 TCAGCTGCTTGTGCTGCCACTTATCCGCTTTCGCTTCTGCTGTTC 284
Db 52 ACAS-----TGCAGCGCGCCACTTGTTCGCGCTCTGCGTCTCC 16

RESULT 13
US-09-513-999C-14808/c
; Sequence 14808, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 14808
; LENGTH: 602
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 8
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 9
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 266
; OTHER INFORMATION: k=q or t

```

; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 553
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 554
; OTHER INFORMATION: k=g or t
US-09-513-999C-14809

Query Match 1.6%; Score 100.8; DB 4; Length 602;
Best Local Similarity 86.7%; Pred. No. 1.4e-18;
Matches 111; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
Qy 123 AAAGCTTATCATCTGCTTCCACTTATCCAAAGCTGCTATGCCCACTGCTGTGCGG 182
Db |||||||
259 ACAGGCTTATCATCTATCTCCACTTATCCAAAGCGCTACGGCCACTTCTGTGCGG 200
Qy 183 CACCTGGAACCAACCGCCAGCCCACTACTGCTCCCACTACCACTGGTTCTCCCACTGA 242
Db |||||||
199 CAGTTGGAAGACCAACCAAGCCCACTAGAGCTCCGCTACCAACGCTTCTCCGCGCTGT 140
Qy 243 TCAGCTGC 250
Db |||||||
139 ACAGCTGC 132

RESULT 14
US-09-513-999C-14809/c
; Sequence 14809, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; PATENT NO. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 14809
; LENGTH: 692
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 8
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 9
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 266
; OTHER INFORMATION: k=g or t
US-09-513-999C-14809

Query Match 1.6%; Score 100.8; DB 4; Length 692;
Best Local Similarity 86.7%; Pred. No. 1.5e-18;
Matches 111; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
Qy 123 AAAGCTTATCATCTGCTTCCACTTATCCAAAGCTGCTATGCCCACTGCTGTGCGG 182
Db |||||||
259 ACAGGCTTATCATCTATCTCCACTTATCCAAAGCGCTACGGCCACTTCTGTGCGG 200
Qy 183 CACCTGGAACCAACCGCCAGCCCACTACTGCTCCCACTACCACTGGTTCTCCCACTGA 242
Db |||||||
199 CAGTTGGAAGACCAACCAAGCCCACTAGAGCTCCGCTACCAACGCTTCTCCGCGCTGT 140

Qy 243 TCAGCTGC 250
Db |||||||
139 ACAGCTGC 132
RESULT 15
US-09-513-999C-14947/c
; Sequence 14947, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; PATENT NO. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 14947
; LENGTH: 842
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 8
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 9
; OTHER INFORMATION: n=a, g, c or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 266
; OTHER INFORMATION: k=g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 436
; OTHER INFORMATION: n=a, g, c or t
US-09-513-999C-14947

Query Match 1.6%; Score 100.8; DB 4; Length 842;
Best Local Similarity 86.7%; Pred. No. 1.8e-18;
Matches 111; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
Qy 123 AAAGCTTATCATCTGCTTCCACTTATCCAAAGCTGCTATGCCCACTGCTGTGCGG 182
Db |||||||
259 ACAGGCTTATCATCTATCTCCACTTATCCAAAGCGCTACGGCCACTTCTGTGCGG 200
Qy 183 CACCTGGAACCAACCGCCAGCCCACTACTGCTCCCACTACCACTGGTTCTCCCACTGA 242
Db |||||||
199 CAGTTGGAAGACCAACCAAGCCCACTAGAGCTCCGCTACCAACGCTTCTCCGCGCTGT 140
Qy 243 TCAGCTGC 250
Db |||||||
139 ACAGCTGC 132

Search completed: July 9, 2005, 08:02:28
Job time : 1015 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 9, 2005, 02:53:40 ; Search time 3333 Seconds
(without alignments)
11882.270 Million cell updates/sec

Title: US-10-758-672A-1
Perfect score: 6308
Sequence: 1 gccaaagattcgacgagg.....aattttgtattgtgtttt 6308

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 6330943 seqs, 3139157217 residues

Total number of hits satisfying chosen parameters: 12661886

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
23: /cgn2_6/ptodata/2/pubpna/US11B_PUBCOMB.seq.*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6308	100.0	6308	19	US-10-758-672A-1
2	6308	100.0	6308	21	US-10-758-636A-1
3	5592	88.6	7742	16	US-10-287-218-22
4	5592	88.6	7742	19	US-10-474-291-22
5	5466	86.7	5466	20	US-10-357-819-9
6	5136	81.4	5205	21	US-10-758-672A-18
7	5136	81.4	5205	21	US-10-758-636A-18

Sequence 1, Appli
Sequence 1, Appli
Sequence 22, Appli
Sequence 22, Appli
Sequence 9, Appli
Sequence 18, Appli
Sequence 18, Appli

8	3055.8	48.4	3059	17	US-10-094-749-916	Sequence 916, App
9	1381.8	21.9	6300	19	US-10-758-672A-3	Sequence 3, Appli
10	1381.8	21.9	6300	21	US-10-758-636A-3	Sequence 3, Appli
11	1233.6	19.6	6089	19	US-10-758-672A-5	Sequence 5, Appli
12	1233.6	19.6	6089	21	US-10-758-636A-5	Sequence 5, Appli
13	981	15.6	6158	19	US-10-755-889-634	Sequence 634, App
14	813.6	12.9	3327	9	US-09-529-063-56	Sequence 56, Appl
15	813.6	12.9	3327	16	US-10-414-378-56	Sequence 56, Appl
16	813.6	12.9	3502	9	US-09-529-063-57	Sequence 57, Appl
17	813.6	12.9	3502	16	US-10-414-378-57	Sequence 57, Appl
18	438.4	6.9	505	10	US-09-918-995-36721	Sequence 36721, A
19	374.8	5.9	1772	17	US-10-264-237-1139	Sequence 1139, Ap
20	250.6	4.0	972	9	US-09-822-849A-532	Sequence 532, App
21	219.8	3.5	4573	13	US-10-071-766-7	Sequence 7, Appli
22	212.2	3.4	578	13	US-10-027-632-260533	Sequence 260533,
23	212.2	3.4	578	17	US-10-027-632-260533	Sequence 260533,
24	160.2	2.5	3411	17	US-10-104-047-183	Sequence 183, App
25	155.6	2.5	469	17	US-10-242-535A-26847	Sequence 26847, A
26	155.6	2.5	469	18	US-10-085-783A-26847	Sequence 26847, A
27	142	2.3	643	17	US-10-264-237-108	Sequence 108, App
28	136.6	2.2	550	16	US-10-029-386-12897	Sequence 12897, A
29	117	1.9	121	16	US-10-029-386-26597	Sequence 26597, A
30	116.2	1.8	364	17	US-10-242-535A-21996	Sequence 21996, A
31	116.2	1.8	364	18	US-10-085-783A-21996	Sequence 21996, A
32	105.6	1.7	1459	17	US-10-264-049-746	Sequence 746, App
33	105.6	1.7	1461	9	US-09-925-301-490	Sequence 490, App
34	105.2	1.7	1692	20	US-10-335-053-82	Sequence 82, Appl
35	104.8	1.7	507	20	US-10-357-930-59123	Sequence 59123, A
36	100.8	1.6	697	9	US-09-823-245A-172	Sequence 172, App
37	100.8	1.6	1370	19	US-10-775-169-131	Sequence 131, App
38	100.8	1.6	1370	20	US-10-684-422-154	Sequence 154, App
39	100.8	1.6	1370	21	US-10-956-157-1949	Sequence 1949, Ap
40	100.8	1.6	1370	21	US-10-956-157-7184	Sequence 7184, App
41	100.6	1.6	1382	14	US-10-016-349A-118	Sequence 118, App
42	96	1.5	436	21	US-10-278-698-446	Sequence 446, App
43	96	1.5	436	21	US-10-278-698-960	Sequence 960, App
44	95.2	1.5	516	21	US-10-278-698-457	Sequence 457, App
45	95.2	1.5	516	21	US-10-278-698-972	Sequence 972, App

ALIGNMENTS

RESULT 1
US-10-758-672A-1
; Sequence 1, Application US/10758672A
; Publication No. US20040185037A1
; GENERAL INFORMATION:
; APPLICANT: Han, et al.
; TITLE OF INVENTION: HUMAN E3 ALPHA UBIQUITIN LIGASE FAMILY
; FILE REFERENCE: 01017/35966B
; CURRENT APPLICATION NUMBER: US/10758.672A
; CURRENT FILING DATE: 2004-01-15
; PRIOR APPLICATION NUMBER: US 09/724,126
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 60/187,911
; PRIOR FILING DATE: 2000-03-08
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 1
; LENGTH: 6308
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (696)..(5942)
US-10-758-672A-1

Query Match 100.0%; Score 6308; DB 19; Length 6308;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCCAAGAATTCGGCAGGGGAAAGCTGAGCCAGAACCAATTAATTCTTGTACTTCTTACCT 60

[illegible]

Db	1081	AGGACGTGTTCTATAAAAATCATCGTTTACAAGATGCATCTTCTACTCGGAGGAGGTTC	1144
Qy	1141	GTGACTGTGGAGACACAGAGCGCATGGAATACTGGCCCTTTTTTGTGTAAATCATGAACCTG	1200
Db	1141	GTGACTGTGGAGACACAGAGCGCATGGAATACTGGCCCTTTTTTGTGTAAATCATGAACCTG	1200
Qy	1201	GAAGACAGGTACTATATAAAGAGAAATTCACGCCTGTCGGTTGAATGAAGAGTTAATTGTCC	1260
Db	1201	GAAGACAGGTACTATATAAAGAGAAATTCACGCCTGTCGGTTGAATGAAGAGTTAATTGTCC	1260
Qy	1261	AAGCCAGGAAAAATATTTCCITTCAGTGCATAAAAATATGTCGTAGAAAACTATATATGGGAAG	1320
Db	1261	AAGCCAGGAAAAATATTTCCITTCAGTGCATAAAAATATGTCGTAGAAAACTATATATGGGAAG	1320
Qy	1321	AGGAAAAAGAACTGCCTCTCTGAACTCCAGATAAGGGAGAAAAATGAAAGATACTATTGTG	1380
Db	1321	AGGAAAAAGAACTGCCTCTCTGAACTCCAGATAAGGGAGAAAAATGAAAGATACTATTGTG	1380
Qy	1381	TCCTTTTCAATGATGAACAACAATTCATATGACACACGTCATATACAGCCTACAAGAGCTC	1440
Db	1381	TCCTTTTCAATGATGAACAACAATTCATATGACACACGTCATATACAGCCTACAAGAGCTC	1440
Qy	1441	TTGACTGTGAGCTCGCAGAGGCCAGTGTGCATACCACTGCCATTCGACAAAGAGGCTCGTC	1500
Db	1441	TTGACTGTGAGCTCGCAGAGGCCAGTGTGCATACCACTGCCATTCGACAAAGAGGCTCGTC	1500
Qy	1501	GGGCTGTTAAAGCGGGAGCTTATGCTGCTTGCACGAAGCAAAAGGAAGATATAAAGAGTC	1560
Db	1501	GGGCTGTTAAAGCGGGAGCTTATGCTGCTTGCACGAAGCAAAAGGAAGATATAAAGAGTC	1560
Qy	1561	ATTCAGAAAAATGTCCTCAACATCCACTTCATGTAGAAGTATTACACTCAGAGATTATGG	1620
Db	1561	ATTCAGAAAAATGTCCTCAACATCCACTTCATGTAGAAGTATTACACTCAGAGATTATGG	1620
Qy	1621	CTCATCAGAAATTTGCTTTTCGGCTTTGGTTTCCTGGATGAACAAATTTATCAGCTATTCAA	1680
Db	1621	CTCATCAGAAATTTGCTTTTCGGCTTTGGTTTCCTGGATGAACAAATTTATCAGCTATTCAA	1680
Qy	1681	GTGACTTTTGGCAGATCTTTTGGCCAAAGCATGCTTTAGAGAAAGAACCTGACTCGGAGATC	1740
Db	1681	GTGACTTTTGGCAGATCTTTTGGCCAAAGCATGCTTTAGAGAAAGAACCTGACTCGGAGATC	1740
Qy	1741	CCTGCTCTATAAGCAGGTTAATGCTTTGGGATGCAAGCTTTTATAAAGTGCCTGTAAGA	1800
Db	1741	CCTGCTCTATAAGCAGGTTAATGCTTTGGGATGCAAGCTTTTATAAAGTGCCTGTAAGA	1800
Qy	1801	TCCTTCATGAATTCGATCTTCAGCAGTTTTTTTTTATGGAGATGGATACAAAATACTCTTTG	1860
Db	1801	TCCTTCATGAATTCGATCTTCAGCAGTTTTTTTTTATGGAGATGGATACAAAATACTCTTTG	1860
Qy	1861	CTATGGAATTTGTGAAGTATTATAAACAACTGCAGAAAGAAATATATCAGTGCATCATG	1920
Db	1861	CTATGGAATTTGTGAAGTATTATAAACAACTGCAGAAAGAAATATATCAGTGCATCATG	1920
Qy	1921	ACAGAAGTATCTCTATAACTGCACTTTTCAGTTTCAGATGTTTACTGTTTCTTCTCTGCTC	1980
Db	1921	ACAGAAGTATCTCTATAACTGCACTTTTCAGTTTCAGATGTTTACTGTTTCTTCTCTGCTC	1980
Qy	1981	GACATCTTATGAAGACAGAAATGTTATCTCTGTCAATTCATGAAACTCTCTCTAGAGTTT	2040
Db	1981	GACATCTTATGAAGACAGAAATGTTATCTCTGTCAATTCATGAAACTCTCTCTAGAGTTT	2040
Qy	2041	TACCTGTAGTACTTCGCACAGGAACAATAAAATTCAACTTCCAGGGTTATAGCCAGGACAAAT	2100
Db	2041	TACCTGTAGTACTTCGCACAGGAACAATAAAATTCAACTTCCAGGGTTATAGCCAGGACAAAT	2100
Qy	2101	TGGGAAGAGTATATGCACTTAATATGTGACCTTAAAGTATATCCTGCATCAGCAAAACCCCAA	2160
Db	2101	TGGGAAGAGTATATGCACTTAATATGTGACCTTAAAGTATATCCTGCATCAGCAAAACCCCAA	2160
Qy	2161	TATGGAACGAAAGATTAAAGATGCAAGTTTCCTTGAAGTTTTTCGATCTTTTTTTGAGATTC	2220
Db	2161	TATGGAACGAAAGATTAAAGATGCAAGTTTCCTTGAAGTTTTTCGATCTTTTTTTGAGATTC	2220

QY 2221 TTACTGTATGACAGGATGGAAGAAATCCGAAGACAGGTTGGGCAACACATTGAAAGTG 2280
DB |||||
2221 TTACTGTATGACAGGATGGAAGAAATCCGAAGACAGGTTGGGCAACACATTGAAAGTG 2280
QY 2281 ATCTCTGATTTGGGAGGCTGCCATTTGCTATACAGATGCAATTTGAAGAAATATTTTACTCATGT 2340
DB |||||
2281 ATCTCTGATTTGGGAGGCTGCCATTTGCTATACAGATGCAATTTGAAGAAATATTTTACTCATGT 2340
QY 2341 TCCAAAGAGTGTGCTTGTGTGATGAAGAACTCTTACTTGTGCTTTATAAAGAAATGTCA 2400
DB |||||
2341 TCCAAAGAGTGTGCTTGTGTGATGAAGAACTCTTACTTGTGCTTTATAAAGAAATGTCA 2400
QY 2401 AAGCTGTGATGAGGTGCAAGTCCAGTCTTATCTAGTAGCAAGACAGTAGTACAAATCGT 2460
DB |||||
2401 AAGCTGTGATGAGGTGCAAGTCCAGTCTTATCTAGTAGCAAGACAGTAGTACAAATCGT 2460
QY 2461 GTGGACATAGTTTGGAAACAAAGTCTTACAGAGTATCTGAGGATCTTGTGAACATACATC 2520
DB |||||
2461 GTGGACATAGTTTGGAAACAAAGTCTTACAGAGTATCTGAGGATCTTGTGAACATACATC 2520
QY 2521 TGCCACTCTCTAGGACCCCTTGTGCTCTTCTATGTAAGTCTTAAAGCAGGCTGGGCTGCTGTT 2580
DB |||||
2521 TGCCACTCTCTAGGACCCCTTGTGCTCTTCTATGTAAGTCTTAAAGCAGGCTGGGCTGCTGTT 2580
QY 2581 CAAAGCTGATGAAATTTGTGCTTTTGAAGGACTTTCAAGTAGAGGTACTAGTGGAAATATC 2640
DB |||||
2581 CAAAGCTGATGAAATTTGTGCTTTTGAAGGACTTTCAAGTAGAGGTACTAGTGGAAATATC 2640
QY 2641 CTTTACGTTTCTGTGTTGTTGCTGCCAGTGTGCTGAGATGTCGGAAGAAATGGAC 2700
DB |||||
2641 CTTTACGTTTCTGTGTTGTTGCTGCCAGTGTGCTGAGATGTCGGAAGAAATGGAC 2700
QY 2701 TGTCTCTTATAGCCAGGTGTTTATACCAAGATGTTAAGTGCAGAGAAATGTATG 2760
DB |||||
2701 TGTCTCTTATAGCCAGGTGTTTATACCAAGATGTTAAGTGCAGAGAAATGTATG 2760
QY 2761 ATAAAGATATCATATGCTTCAGATGTTGTCATCTTTAATGATCCCAATAGTCTTGT 2820
DB |||||
2761 ATAAAGATATCATATGCTTCAGATGTTGTCATCTTTAATGATCCCAATAGTCTTGT 2820
QY 2821 TACTGGTACTTCAGAGGTATGAATGTCGAGGCTTTTAAACAGACCATATCTACAAAG 2880
DB |||||
2821 TACTGGTACTTCAGAGGTATGAATGTCGAGGCTTTTAAACAGACCATATCTACAAAG 2880
QY 2881 ACCAGATTTGATTAACAAATATACTAATAAGAAATGCTTCAGGTCCTCATCT 2940
DB |||||
2881 ACCAGATTTGATTAACAAATATACTAATAAGAAATGCTTCAGGTCCTCATCT 2940
QY 2941 ATATTGGGTGAGGCTTATGTAATGAGTGGGAAATGTGACCAAGAGAGGTCA 3000
DB |||||
2941 ATATTGGGTGAGGCTTATGTAATGAGTGGGAAATGTGACCAAGAGAGGTCA 3000
QY 3001 TGAGAGAAATCAATTCATCTGCTTTCATGATGACCCATGCCACAGTGCCTTGC 3060
DB |||||
3001 TGAGAGAAATCAATTCATCTGCTTTCATGATGACCCATGCCACAGTGCCTTGC 3060
QY 3061 ATTTACCTGAGAAATCAATGAACTGGCTTGAAGAAATGTGATGAAATGAAAGTGGCA 3120
DB |||||
3061 ATTTACCTGAGAAATCAATGAACTGGCTTGAAGAAATGTGATGAAATGAAAGTGGCA 3120
QY 3121 CATTTAAGAAACAGGCTTATCAGGCCATGAGGTTTATGAATCAAAAGATGAAATCACTGA 3180
DB |||||
3121 CATTTAAGAAACAGGCTTATCAGGCCATGAGGTTTATGAATCAAAAGATGAAATCACTGA 3180
QY 3181 AAGACTCAATATGTAATTTTATCTTACTTCAAAACCCAGCATAGGAGGCTGAACATA 3240
DB |||||
3181 AAGACTCAATATGTAATTTTATCTTACTTCAAAACCCAGCATAGGAGGCTGAACATA 3240
QY 3241 TGCAGAGAAAGGAGAAACAAAGAAACAAAGATGAAAGTTCGCGCCACACCCACCTC 3300
DB |||||
3241 TGCAGAGAAAGGAGAAACAAAGAAACAAAGATGAAAGTTCGCGCCACACCCACCTC 3300

QY 3301 CTGAATTTCTGCTGCTTTTTCAGCAAGTGAATTAACCTTCTCAACTGTGATATCATGATGT 3360
DB |||||
3301 CTGAATTTCTGCTGCTTTTTCAGCAAGTGAATTAACCTTCTCAACTGTGATATCATGATGT 3360
QY 3361 ACATTTCTCAGAGCCGCTATTTGAGCGGGCAATAGACACAGATTTCTAACTTGTGGACCGAAG 3420
DB |||||
3361 ACATTTCTCAGAGCCGCTATTTGAGCGGGCAATAGACACAGATTTCTAACTTGTGGACCGAAG 3420
QY 3421 GGATGCTCCAAATGCTTTTTCATATTTCTGGCAATGCGGTTTACTAGAGAAAGCAACAGC 3480
DB |||||
3421 GGATGCTCCAAATGCTTTTTCATATTTCTGGCAATGCGGTTTACTAGAGAAAGCAACAGC 3480
QY 3481 TTCAAAAAGCTCTTGAAGAGAAAGTAAACATTTGATCTTTTATCATAGGCTTCAAGATTTGG 3540
DB |||||
3481 TTCAAAAAGCTCTTGAAGAGAAAGTAAACATTTGATCTTTTATCATAGGCTTCAAGATTTGG 3540
QY 3541 GAAAGTTCTAGCCATGAATATACAAATGCTTTTGGAAAGAACTCAAGAGAAATTTCCCCAGTTAG 3600
DB |||||
3541 GAAAGTTCTAGCCATGAATATACAAATGCTTTTGGAAAGAACTCAAGAGAAATTTCCCCAGTTAG 3600
QY 3601 AAGGCCAAGAGCAATGATTAACGTTGATCTTCAAGATGTTTGACACAGTGAAGCGATTA 3660
DB |||||
3601 AAGGCCAAGAGCAATGATTAACGTTGATCTTCAAGATGTTTGACACAGTGAAGCGATTA 3660
QY 3661 GAGAAAAATCTTGTTTAATTTCTAGCAACCAATCAGGATCGGAATCTATTAAAGATGATG 3720
DB |||||
3661 GAGAAAAATCTTGTTTAATTTAGCAACCAATCAGGATCGGAATCTATTAAAGATGATG 3720
QY 3721 AGATTTACTCATGATAAGAAAGCAAGAAAGCAAGAAAGCAAGCTGAAGCTGCTAGGCTAC 3780
DB |||||
3721 AGATTTACTCATGATAAGAAAGCAAGAAAGCAAGAAAGCAAGCTGCTAGGCTAC 3780
QY 3781 ATGCGCAGAGATCATGCTCAGATGCTGCTGCTTACAGAAAACTTCAATTTGAAATCTCAT 3840
DB |||||
3781 ATGCGCAGAGATCATGCTCAGATGCTGCTGCTTACAGAAAACTTCAATTTGAAATCTCAT 3840
QY 3841 AACTCATGTATGACATACATCAGAAATGCTGCGGAAAGAAAGATTTCCATTATGGAGGAG 3900
DB |||||
3841 AACTCATGTATGACATACATCAGAAATGCTGCGGAAAGAAAGATTTCCATTATGGAGGAG 3900
QY 3901 AGAGCAACCCAGCAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3960
DB |||||
3901 AGAGCAACCCAGCAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3960
QY 3961 CTGTTACTGAAAAGAGGAGTGTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4020
DB |||||
3961 CTGTTACTGAAAAGAGGAGTGTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4020
QY 4021 TAGAAAAATTAATGCGCATTTATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4080
DB |||||
4021 TAGAAAAATTAATGCGCATTTATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4080
QY 4081 ACAGGGGAAACCCATAGAACTCTCAGGAGAGCCCTAGAGCCCTTTTTCATGGATCCAG 4140
DB |||||
4081 ACAGGGGAAACCCATAGAACTCTCAGGAGAGCCCTAGAGCCCTTTTTCATGGATCCAG 4140
QY 4141 ACTTGGCATATGGAACCTTATACAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4200
DB |||||
4141 ACTTGGCATATGGAACCTTATACAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4200
QY 4201 AGAAGTATTTTGAAGCTGTAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4260
DB |||||
4201 AGAAGTATTTTGAAGCTGTAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4260
QY 4261 ACTTGGAAAGTGGAGAAATATCTTTGCGCTCTTTGCAAACTCTGCTGCAATCTGTGATCC 4320
DB |||||
4261 ACTTGGAAAGTGGAGAAATATCTTTGCGCTCTTTGCAAACTCTGCTGCAATCTGTGATCC 4320
QY 4321 CCATTTATTTCTTGTCAACCTCAAAAGATAAAGAGTGAAGTGCAGATGCTCTTGTCTCAAC 4380
DB |||||
4321 CCATTTATTTCTTGTCAACCTCAAAAGATAAAGAGTGAAGTGCAGATGCTCTTGTCTCAAC 4380
QY 4381 TTTTGAACCTTGGCAGCGGTGGGATACAGACTGTTTCTGGCCAGAAATATCAGGTTATATATA 4440

Db 4381 TTTTGACCTGGCAGCGTGGATACAGACTGTTCTGGCCAGAAATATCAGGTATATATAA 4440
Qy 4441 GACATGCTAAAGGAGAAAACCAATTCCTATTTCTTTTAAATCAAGGAATGGAGATCTTA 4500
Db 4441 GACATGCTAAAGGAGAAAACCAATTCCTATTTCTTTTAAATCAAGGAATGGAGATCTTA 4500
Qy 4501 CTTTGGAGTTCCATTCGATCCTGAGTTTGGGGTTGAGTCTTCGATTAATAATTCAAATA 4560
Db 4501 CTTTGGAGTTCCATTCGATCCTGAGTTTGGGGTTGAGTCTTCGATTAATAATTCAAATA 4560
Qy 4561 GCATCAAGGAAATGGTTATTTCTCTTGGCCACAACAATTTATAGAAATTCGATTTGAAAGTGC 4620
Db 4561 GCATCAAGGAAATGGTTATTTCTCTTGGCCACAACAATTTATAGAAATTCGATTTGAAAGTGC 4620
Qy 4621 CACCTGATGAAGGGATCCTCGAGTCCCATGCTGACCTGGAGCACCTGCGCTTTCACTA 4680
Db 4621 CACCTGATGAAGGGATCCTCGAGTCCCATGCTGACCTGGAGCACCTGCGCTTTCACTA 4680
Qy 4681 TCCAGGCAATTCGAATCTATTTGGGAGATGAAGGAACCTCTGTTTGGAGCACTTCAAA 4740
Db 4681 TCCAGGCAATTCGAATCTATTTGGGAGATGAAGGAACCTCTGTTTGGAGCACTTCAAA 4740
Qy 4741 ATAGCAGCAATAATGGTCTGAAGCAATTAATGTCAGTTTGCAGATTCGACAGGAGTTACCT 4800
Db 4741 ATAGCAGCAATAATGGTCTGAAGCAATTAATGTCAGTTTGCAGATTCGACAGGAGTTACCT 4800
Qy 4801 GTCCTCAGGTCCTGTATACAGAAACATCTGGTTCCTTCTATCATAGTTGTTCTTCTTAACA 4860
Db 4801 GTCCTCAGGTCCTGTATACAGAAACATCTGGTTCCTTCTATCATAGTTGTTCTTCTTAACA 4860
Qy 4861 TAAATACAGAGATACACATGCTTCTGTATGGGATGACCTGTTGATCTGCAGGCTTCTTCAG 4920
Db 4861 TAAATACAGAGATACACATGCTTCTGTATGGGATGACCTGTTGATCTGCAGGCTTCTTCAG 4920
Qy 4921 CTGTGTAGCATTCCTTCTGTATGGGATGACCTGTTGATCTGCAGGCTTCTTCAG 4980
Db 4921 CTGTGTAGCATTCCTTCTGTATGGGATGACCTGTTGATCTGCAGGCTTCTTCAG 4980
Qy 4981 TTAGTTCTTCTTATACCAACCTTTATCTCTCCATTTGATCACCAGTGGCAACATGCTTC 5040
Db 4981 TTAGTTCTTCTTATACCAACCTTTATCTCTCCATTTGATCACCAGTGGCAACATGCTTC 5040
Qy 5041 AGATCTACTTACAGTAGACACAGGCTTACCGCTTGTCTCAGTTTCAAGAGACAGTGAAG 5100
Db 5041 AGATCTACTTACAGTAGACACAGGCTTACCGCTTGTCTCAGTTTCAAGAGACAGTGAAG 5100
Qy 5101 AGGCTCATTTCCGATCTTCTTCTTTCAGAAATTTCTCAATATACAAAGTGGCTCCATTG 5160
Db 5101 AGGCTCATTTCCGATCTTCTTCTTTCAGAAATTTCTCAATATACAAAGTGGCTCCATTG 5160
Qy 5161 GGTGTGATATTCCTGGCTGGTATTTGTGGGTCTCACTGAAGAAATGGCAATCCACCTTATC 5220
Db 5161 GGTGTGATATTCCTGGCTGGTATTTGTGGGTCTCACTGAAGAAATGGCAATCCACCTTATC 5220
Qy 5221 TTCGCTGTGCTGATTTTCTTCCATTAATTTTGGGTAATCTCGGCTGAGGAACCTGC 5280
Db 5221 TTCGCTGTGCTGATTTTCTTCCATTAATTTTGGGTAATCTCGGCTGAGGAACCTGC 5280
Qy 5281 ATACCATAATTCGAGAGAGAGTACAGTGCATCTGTAGCTATCTATCTTTACCTACAA 5340
Db 5281 ATACCATAATTCGAGAGAGAGTACAGTGCATCTGTAGCTATCTATCTTTACCTACAA 5340
Qy 5341 ATTTGTCTCTCTTCTTCCAGGAATATTTGGATCTGTAAGGCCCTTGTCTCCAGAGGTGGT 5400
Db 5341 ATTTGTCTCTCTTCTTCCAGGAATATTTGGATCTGTAAGGCCCTTGTCTCCAGAGGTGGT 5400
Qy 5401 GTGAGATCTCGCTTCTTCCAGGAATATTTGGATCTGTAAGGCCCTTGTCTCCAGAGGTGGT 5460
Db 5401 GTGAGATCTCGCTTCTTCCAGGAATATTTGGATCTGTAAGGCCCTTGTCTCCAGAGGTGGT 5460
Qy 5461 GAAAGAAATAGTTTGTATAGAGCTTCTCTGATGACTATATAGTCTGCTCTGTAATCAAGCTT 5520

Db 5461 GAAAGAAATAGTTTGTATAGAGCTTCTCTGATGACTATAGTCTGCTCTGTAATCAAGCTT 5520
Qy 5521 CTCAATTTTCCAGTCCCAAGCTCTGCGATGATGAGCGAAGACATCTCTGCTCTCTGCTTTT 5580
Db 5521 CTCAATTTTCCAGTCCCAAGCTCTGCGATGATGAGCGAAGACATCTCTGCTCTCTGCTTTT 5580
Qy 5581 TCTGTGGGGCTTACTATATTTCTCAGAAACATTTTCTGCCAGGAAATTTGTGAACGGGAAG 5640
Db 5581 TCTGTGGGGCTTACTATATTTCTCAGAAACATTTTCTGCCAGGAAATTTGTGAACGGGAAG 5640
Qy 5641 AGGTGTGAGCTTGCAATTTTTCACGCACTTCACTGTGAGCGGAGTCTGCAATTTTCTTAA 5700
Db 5641 AGGTGTGAGCTTGCAATTTTTCACGCACTTCACTGTGAGCGGAGTCTGCAATTTTCTTAA 5700
Qy 5701 AAATCAGAGATGCGGAGTCTGCTGTTGAAGTAAAGCCAGAGGCTGTGCCATTCAG 5760
Db 5701 AAATCAGAGATGCGGAGTCTGCTGTTGAAGTAAAGCCAGAGGCTGTGCCATTCAG 5760
Qy 5761 CTCCTTACTTGGATGAATATGGGAAACAGACCTGCTGAAAGAGGGGCAACCCCTTC 5820
Db 5761 CTCCTTACTTGGATGAATATGGGAAACAGACCTGCTGAAAGAGGGGCAACCCCTTC 5820
Qy 5821 ATTTATCTGTGAGCGGTATCGGAAGCTTCAATTTTGGTCTGGCAACAACATCTGCAATTAAG 5880
Db 5821 ATTTATCTGTGAGCGGTATCGGAAGCTTCAATTTTGGTCTGGCAACAACATCTGCAATTAAG 5880
Qy 5881 AAGGATGCTTGAAGCCAGAGAGTAAATCAGATGTTTATTTGGATTCACCTGGCAGTTAC 5940
Db 5881 AAGGATGCTTGAAGCCAGAGAGTAAATCAGATGTTTATTTGGATTCACCTGGCAGTTAC 5940
Qy 5941 TGTGAGCTTCAACTCTGCTCAAGCAACATCAAAATGACGACAGTAAAGGCTGATTC 6000
Db 5941 TGTGAGCTTCAACTCTGCTCAAGCAACATCAAAATGACGACAGTAAAGGCTGATTC 6000
Qy 6001 AAAATTAAGGAAACTTTCTGAGGCTGGGAAAGTATTTGGAGGCTCTTTTCTCATGTC 6060
Db 6001 AAAATTAAGGAAACTTTCTGAGGCTGGGAAAGTATTTGGAGGCTCTTTTCTCATGTC 6060
Qy 6061 CAGGTTCACTTACATCAATTAATTTCTTAAAGGATGTTCTTCAATTAGCAACA 6120
Db 6061 CAGGTTCACTTACATCAATTAATTTCTTAAAGGATGTTCTTCAATTAGCAACA 6120
Qy 6121 TATGCTTCAAGGAAAAAGGACATAGATCAATCTGTTTATGTGCTAGTATTTCCAGGA 6180
Db 6121 TATGCTTCAAGGAAAAAGGACATAGATCAATCTGTTTATGTGCTAGTATTTCCAGGA 6180
Qy 6181 ATTTATTTCCCTTCAATTTTGTCTCAATTTTATTTTATTTTCACTGCTGATGAG 6240
Db 6181 ATTTATTTCCCTTCAATTTTGTCTCAATTTTATTTTATTTTCACTGCTGATGAG 6240
Qy 6241 TCAGTCAAAACAGTTGATGATTTTATGTTGTTAACTCTTCTGCAATTTTGTATTT 6300
Db 6241 TCAGTCAAAACAGTTGATGATTTTATGTTGTTAACTCTTCTGCAATTTTGTATTT 6300
Qy 6301 GGTGTTTTT 6308
Db 6301 GGTGTTTTT 6308

RESULT 2
US-10-758-636A-1
; Sequence 1, Application US/10758636A
; Publication No. US20050089876A1
; GENERAL INFORMATION:
; APPLICANT: Han, et al.
; TITLE OF INVENTION: HUMAN E3 ALPHA UBIQUITIN LIGASE FAMILY
; FILE REFERENCE: 01017/35966C
; CURRENT APPLICATION NUMBER: US/10758, 636A
; CURRENT FILING DATE: 2004-01-15
; PRIOR APPLICATION NUMBER: US 09/724,126
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 60/187,911
; PRIOR FILING DATE: 2000-03-08


```
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 6308
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (696)..(5942)
US-10-758-636A-1

Query Match      100.0%; Score 6308; DB 21; Length 6308;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCAAGAATTCCGCGCAGGGGAAAGAGCTGAGCCAGGAACCAAAATTAATCTGCTTTTACCT 60
DB 1 GCCAAGAATTCCGCGCAGGGGAAAGAGCTGAGCCAGGAACCAAAATTAATCTGCTTTTACCT 60

QY 61 CATTTGTGAAGCAAGCGTCAAAAACAGCTTCAACTATCTTGAACAAGAGAACTTACCT 120
DB 61 CATTTGTGAAGCAAGCGTCAAAAACAGCTTCAACTATCTTGAACAAGAGAACTTACCT 120

QY 121 CCAAAGGCTTATCATCTGCTTCCACTTATCCAAAGCTGCTATGCGCACTGCGCTGTC 180
DB 121 CCAAAGGCTTATCATCTGCTTCCACTTATCCAAAGCTGCTATGCGCACTGCGCTGTC 180

QY 181 CGCACTTGGAAACACCGCCAGCCCACTACTGCTCCCACTACCACTGCTTCTCCCACTCT 240
DB 181 CGCACTTGGAAACACCGCCAGCCCACTACTGCTCCCACTACCACTGCTTCTCCCACTCT 240

QY 241 GATCAGCTGCTTGTGCTGCCATCTATTCGGCTTCTGCTGCTTCTGAGTAAATGATACA 300
DB 241 GATCAGCTGCTTGTGCTGCCATCTATTCGGCTTCTGCTGCTTCTGAGTAAATGATACA 300

QY 301 CACCTGTGAACACCACTTACTTCTGCTGCTATGATTTGACTCTAGCTGGATC 360
DB 301 CACCTGTGAACACCACTTACTTCTGCTGCTATGATTTGACTCTAGCTGGATC 360

QY 361 CCGAGCTTTTGTACACATGTGCAAGTGCCTCCAGGGGTAGAACTCTTAAATAAGAGAT 420
DB 361 CCGAGCTTTTGTACACATGTGCAAGTGCCTCCAGGGGTAGAACTCTTAAATAAGAGAT 420

QY 421 GTATGCAACAGTTCCAGCACCAAAACCCAGATATACCACTTACAGTACCAAGAGCTAC 480
DB 421 GTATGCAACAGTTCCAGCACCAAAACCCAGATATACCACTTACAGTACCAAGAGCTAC 480

QY 481 GCCTGATAAATTAGAGGGGAAAAAATAATCTCCAGTCCCTTCACTGCTGAGCGTTGCT 540
DB 481 GCCTGATAAATTAGAGGGGAAAAAATAATCTCCAGTCCCTTCACTGCTGAGCGTTGCT 540

QY 541 TCCGGGAAGCGGCGGGAAGCACTCTCCGAGTCTGCGTCAAAACCCGACTTTCAGGGGCG 600
DB 541 TCCGGGAAGCGGCGGGAAGCACTCTCCGAGTCTGCGTCAAAACCCGACTTTCAGGGGCG 600

QY 601 TCGTAAAAAGTGTGCTCCCTGCTCTCCGACCGGCCACAGGTTTCGCTTTCGCTGCGCG 660
DB 601 TCGTAAAAAGTGTGCTCCCTGCTCTCCGACCGGCCACAGGTTTCGCTTTCGCTGCGCG 660

QY 661 GGGTTCGCAACTGACAGCGTCACTTTCCTCAAGATGCGGACAGAGGCGTGGAGGTA 720
DB 661 GGGTTCGCAACTGACAGCGTCACTTTCCTCAAGATGCGGACAGAGGCGTGGAGGTA 720

QY 721 CTGAGAGGATGGAATCAGCGGGAGTTTACCCAGACCCCTCAGCGTCTGGCATCTTGGT 780
DB 721 CTGAGAGGATGGAATCAGCGGGAGTTTACCCAGACCCCTCAGCGTCTGGCATCTTGGT 780

QY 781 GGGATCAGCAAGTTGATTTTATATCTGCTTCTTGCATCATTTGGCAAAATTTGGTCCAG 840
DB 781 GGGATCAGCAAGTTGATTTTATATCTGCTTCTTGCATCATTTGGCAAAATTTGGTCCAG 840

QY 841 AAATTTACTTGTGTAATGGAACCCAGACTTGGAAAAAGCAGAGGAAAGTGTACAAATGT 900
DB 841 AAATTTACTTGTGTAATGGAACCCAGACTTGGAAAAAGCAGAGGAAAGTGTACAAATGT 900
```

```
DB 841 AAATTTACTTGTGTAATGGAACCCAGACTTGGAAAAAGCAGAGGAAAGTGTACAAATGT 900
QY 901 CAATATTTCACTCCACTCGGATGTTACTTATTTTGGAGAGATCCAGATATTTCTCTAGAGA 960
DB 901 CAATATTTCACTCCACTCGGATGTTACTTATTTTGGAGAGATCCAGATATTTCTCTAGAGA 960
QY 961 AATTGAAGCAGAGTGGAGCATTTTCACTTTTGTGGAGGGTTTTTCAAAAGTGGAGAGACAA 1020
DB 961 AATTGAAGCAGAGTGGAGCATTTTCACTTTTGTGGAGGGTTTTTCAAAAGTGGAGAGACAA 1020
QY 1021 CCTATTCTTGCAGGATTTGTCAATTTGATCCAAATGTCATCTCTGTCATGTCCTCTCC 1080
DB 1021 CCTATTCTTGCAGGATTTGTCAATTTGATCCAAATGTCATCTCTGTCATGTCCTCTCC 1080
QY 1081 AGGACAGTGTTCATAAAAATCATCGTTTACAGATGCATATCTTCTACTTGGAGAGGTTCT 1140
DB 1081 AGGACAGTGTTCATAAAAATCATCGTTTACAGATGCATATCTTCTACTTGGAGAGGTTCT 1140
QY 1141 GTGACTGTGGAGACACAGAGGATGGAACCTGGCCCTTTTGTGTAAATCATGAACCTG 1200
DB 1141 GTGACTGTGGAGACACAGAGGATGGAACCTGGCCCTTTTGTGTAAATCATGAACCTG 1200
QY 1201 GAAGAGCAGGTACTATTAAGAGAAATTCACGCTGTCGTTGAATCAAGAGGTAATTTGTCC 1260
DB 1201 GAAGAGCAGGTACTATTAAGAGAAATTCACGCTGTCGTTGAATCAAGAGGTAATTTGTCC 1260
QY 1261 AAGCAGGAAATATTTCTTCTAGTGAATAATATGTCGTAGAAATGACTATATGCGAAG 1320
DB 1261 AAGCAGGAAATATTTCTTCTAGTGAATAATATGTCGTAGAAATGACTATATGCGAAG 1320
QY 1321 AGGAAAAAGAACTGCTCTCTGAACTCCAGATAAGGAGAGAAATAAGAAAGATATCTATTGTG 1380
DB 1321 AGGAAAAAGAACTGCTCTCTGAACTCCAGATAAGGAGAGAAATAAGAAAGATATCTATTGTG 1380
QY 1381 TCCTTTTCAATGATGAAACCACTTATCATATGACCACTATATACAGCCCTACAAAGAGCTC 1440
DB 1381 TCCTTTTCAATGATGAAACCACTTATCATATGACCACTATATACAGCCCTACAAAGAGCTC 1440
QY 1441 TTGACTGTGAGCTCGCAGAGGCCCACTGTCATACCACTGCCATTCGCAAAAGAGGTCGTC 1500
DB 1441 TTGACTGTGAGCTCGCAGAGGCCCACTGTCATACCACTGCCATTCGCAAAAGAGGTCGTC 1500
QY 1501 GGGCTGTTAAAGCGGGAGCTTATGCTGCTTGCAGGAAAGCAAGAGATATAAAGAGTTC 1560
DB 1501 GGGCTGTTAAAGCGGGAGCTTATGCTGCTTGCAGGAAAGCAAGAGATATAAAGAGTTC 1560
QY 1561 ATTCAGAAAAATGCTCTCAACATCCATTCATGATGAGTATACCTCAGAGATTTATGG 1620
DB 1561 ATTCAGAAAAATGCTCTCAACATCCATTCATGATGAGTATACCTCAGAGATTTATGG 1620
QY 1621 CTCATCAAAAATTTGCTTTGCTGCTTGGTCTCTGGATGAAACAAAATTTATGAGCTATTCAA 1680
DB 1621 CTCATCAAAAATTTGCTTTGCTGCTTGGTCTCTGGATGAAACAAAATTTATGAGCTATTCAA 1680
QY 1681 GTGACTTTAGCAGATCTTTTGCAGCATGCCCTTAGAGAGAAACCTGACTCGGAGATC 1740
DB 1681 GTGACTTTAGCAGATCTTTTGCAGCATGCCCTTAGAGAGAAACCTGACTCGGAGATC 1740
QY 1741 CCTGCTCATGAAGCAGGTTAATGCTTTGGGATGCAAGCTTTTATAAGGTCGCCGTAGA 1800
DB 1741 CCTGCTCATGAAGCAGGTTAATGCTTTGGGATGCAAGCTTTTATAAGGTCGCCGTAGA 1800
QY 1801 TCCTTCTGAATTTGATCTTTCAGCAGTTTCTTATGAGAGTGAATACAAAACCTCTTTG 1860
DB 1801 TCCTTCTGAATTTGATCTTTCAGCAGTTTCTTATGAGAGTGAATACAAAACCTCTTTG 1860
QY 1861 CTATGGAATTTGTGAAGTATTAACCACTGCAAGAGAAATATATCATGATGATCATG 1920
DB 1861 CTATGGAATTTGTGAAGTATTAACCACTGCAAGAGAAATATATCATGATGATCATG 1920
QY 1921 ACAGAGTATCTCTATTAACCTGCACTTTTCAGTTTCAGATGTTTCTTCTACTCTGGCTC 1980
DB 1921 ACAGAGTATCTCTATTAACCTGCACTTTTCAGTTTCAGATGTTTCTTCTACTCTGGCTC 1980
```


|||||
Db 4141 ACTTGGCATATGGAACTTATACAGAACTGTGGTCATGTAATGACGAGTGTCTGGC 4200
Qy 4201 AGAAGTATTTGAAGCTGTACAGCTGAGCTCTCAGCAGCGCATTCATGTTGACCTTTTGG 4260
Db 4201 AGAAGTATTTGAAGCTGTACAGCTGAGCTCTCAGCAGCGCATTCATGTTGACCTTTTGG 4260
Qy 4261 ACTTGGAAAGTGGAGAAATATCTTTGCCCTCTTTGGCAATCTCTGTGCAATACTGTGATCC 4320
Db 4261 ACTTGGAAAGTGGAGAAATATCTTTGCCCTCTTTGGCAATCTCTGTGCAATACTGTGATCC 4320
Qy 4321 CCAATTATCTTTGCAACTCAAAAGATAAACAAGTGAAGTGCAGATGCTCTGTGCTCAAC 4380
Db 4321 CCAATTATCTTTGCAACTCAAAAGATAAACAAGTGAAGTGCAGATGCTCTGTGCTCAAC 4380
Qy 4381 TTTTGACCTCTGCGAGTGGATACAGACTGTTCTGGCCAGAAATACAGGTTATATATAA 4440
Db 4381 TTTTGACCTCTGCGAGTGGATACAGACTGTTCTGGCCAGAAATACAGGTTATATATAA 4440
Qy 4441 GACATGCTAAAGGAGAAACCAATCTCTTCTTTAAATCAAGGAATGGAGATTTCTA 4500
Db 4441 GACATGCTAAAGGAGAAACCAATCTCTTCTTTAAATCAAGGAATGGAGATTTCTA 4500
Qy 4501 CTTTGGAGTCCATTCCTGATCTGAGTCTTGGGTTGAGTCTTGGATTAATTAATA 4560
Db 4501 CTTTGGAGTCCATTCCTGATCTGAGTCTTGGGTTGAGTCTTGGATTAATTAATA 4560
Qy 4561 GCATCAAGGAATGGTATTTCTTTGGCCAAACAATTTATAGAAATGGATTTGAAAGTGC 4620
Db 4561 GCATCAAGGAATGGTATTTCTTTGGCCAAACAATTTATAGAAATGGATTTGAAAGTGC 4620
Qy 4621 CACCTGATGAAGGGATCCTGAGTCCCATCTGACCTGAGGACCTGCGCTTTCACTA 4680
Db 4621 CACCTGATGAAGGGATCCTGAGTCCCATCTGACCTGAGGACCTGCGCTTTCACTA 4680
Qy 4681 TCCAGGCAATGGAATCTATTTGGAGATGAAGGAACCTCTGTTGGAGCACTTCAAA 4740
Db 4681 TCCAGGCAATGGAATCTATTTGGAGATGAAGGAACCTCTGTTGGAGCACTTCAAA 4740
Qy 4741 ATAGCAGCATTAATGGTCTGAAGCATTAATGTCAGTCTGAGTTCACAGAGGATTACCT 4800
Db 4741 ATAGCAGCATTAATGGTCTGAAGCATTAATGTCAGTCTGAGTTCACAGAGGATTACCT 4800
Qy 4801 GTCTCAGTCTGTATACAGAAACATCTGGTCTTCTATCACTGTTCTTCTCTCAACA 4860
Db 4801 GTCTCAGTCTGTATACAGAAACATCTGGTCTTCTATCACTGTTCTTCTCTCAACA 4860
Qy 4861 TAAATACAGAGATACACATGCTCTGTCTATAGATCTGTTCAATGTTTGGTGGTG 4920
Db 4861 TAAATACAGAGATACACATGCTCTGTCTATAGATCTGTTCAATGTTTGGTGGTG 4920
Qy 4921 CTGTGTTAGCATTCCTCTTATTTGGGATGACCTCTGTTGATCTGAGCTTCTTCAG 4980
Db 4921 CTGTGTTAGCATTCCTCTTATTTGGGATGACCTCTGTTGATCTGAGCTTCTTCAG 4980
Qy 4981 TTAGTCTTCTTATACCAACCTTTATCTCTCCATTTGATCAACATGCGCACATGCTTC 5040
Db 4981 TTAGTCTTCTTATACCAACCTTTATCTCTCCATTTGATCAACATGCGCACATGCTTC 5040
Qy 5041 AGATACTACTTACAGTAGACACAGGCTTACCCCTTGTGCTCAGGTTCAAGAGCAGTGAAG 5100
Db 5041 AGATACTACTTACAGTAGACACAGGCTTACCCCTTGTGCTCAGGTTCAAGAGCAGTGAAG 5100
Qy 5101 AGGCTCATTCGCGATCTTCTTTGAGAAATTTCTCAATATACAGTGGCTCCATTTG 5160
Db 5101 AGGCTCATTCGCGATCTTCTTTGAGAAATTTCTCAATATACAGTGGCTCCATTTG 5160
Qy 5161 GGTGTGATATCTCGGCTGATTTTGGGCTCTCACTGAAGAAATGGCATCCCTTATC 5220
Db 5161 GGTGTGATATCTCGGCTGATTTTGGGCTCTCACTGAAGAAATGGCATCCCTTATC 5220
Qy 5221 TTGCTGTGCTGCAATTTTTCATTTTACATTTTGGGTTAACTCCGCTGAGGAATGTC 5280
|||||

Db 5221 TTGCTGTGCTGCAATTTTTCATTTTACTTTGGGTTAACTCCGCTGAGGAATGTC 5280
Qy 5281 ATACCAATTTGCGAGAGGAGAGTACAGTGCATCTCTGTAGTATCTATCTTTTACTACAA 5340
Db 5281 ATACCAATTTGCGAGAGGAGAGTACAGTGCATCTCTGTAGTATCTATCTTTTACTACAA 5340
Qy 5341 ATTTGTTCTGCTCTTCCNGGAATATTTGGGATCTGTAAAGGCCCTTGTCTCAGAGGTTGT 5400
Db 5341 ATTTGTTCTGCTCTTCCNGGAATATTTGGGATCTGTAAAGGCCCTTGTCTCAGAGGTTGT 5400
Qy 5401 GTGCAGATCTCTGCTCTTAACTGTTTGAAGCAAAAACACCGTGTGTGAGTACCCCTA 5460
Db 5401 GTGCAGATCTCTGCTCTTAACTGTTTGAAGCAAAAACACCGTGTGTGAGTACCCCTA 5460
Qy 5461 GAAAAGAAATAGTCTGATGAGCTTCTGATGACTATAGTGTCTCTGATCAAGCTT 5520
Db 5461 GAAAAGAAATAGTCTGATGAGCTTCTGATGACTATAGTGTCTCTGATCAAGCTT 5520
Qy 5521 CTCAATTTGAGTGTGCGGCTCTGAGATGATGAGCGAAAGCATCTCTGTCTGCTCTT 5580
Db 5521 CTCAATTTGAGTGTGCGGCTCTGAGATGATGAGCGAAAGCATCTCTGTCTGCTCTT 5580
Qy 5581 TCTGTGGGCTTATCTATGTTCTCAGAACTTTGCTGCGAGAAATTTGTGAACGGGGAAG 5640
Db 5581 TCTGTGGGCTTATCTATGTTCTCAGAACTTTGCTGCGAGAAATTTGTGAACGGGGAAG 5640
Qy 5641 AGTTTGGAGCTTGCAATTTTTCACGACTTCACTCTGTGAGCGGAGTCTGCAATTTTCTTAA 5700
Db 5641 AGTTTGGAGCTTGCAATTTTTCACGACTTCACTCTGTGAGCGGAGTCTGCAATTTTCTTAA 5700
Qy 5701 AAATCAGAGATGCGGAGTGTCTCTGTTGAGGTAAAGCCAGAGGCTGTGCTCTATCCAG 5760
Db 5701 AAATCAGAGATGCGGAGTGTCTCTGTTGAGGTAAAGCCAGAGGCTGTGCTCTATCCAG 5760
Qy 5761 CTCTCTTCTGATGATATGAGAAACAGACCTTCTGCTGAGAGGGGCAACCCCTTTC 5820
Db 5761 CTCTCTTCTGATGATATGAGAAACAGACCTTCTGCTGAGAGGGGCAACCCCTTTC 5820
Qy 5821 ATTTATCTGCTGAGCGGTATCGGAAGCTTCAATTTGGTCTGCGCAACACACTGCAATTAAG 5880
Db 5821 ATTTATCTGCTGAGCGGTATCGGAAGCTTCAATTTGGTCTGCGCAACACACTGCAATTAAG 5880
Qy 5881 AAGAGATTTGCTGAGCGGAGAGACTAATCAGATGTTATTTGGATTTCACTGCGAGTTTAC 5940
Db 5881 AAGAGATTTGCTGAGCGGAGAGACTAATCAGATGTTATTTGGATTTCACTGCGAGTTTAC 5940
Qy 5941 TGTGAGCTTCAACTCTGCTCAAGCAATCAAAATGACGACAGTGTAAAGGCTGATTC 6000
Db 5941 TGTGAGCTTCAACTCTGCTCAAGCAATCAAAATGACGACAGTGTAAAGGCTGATTC 6000
Qy 6001 AAAATTTATGAAAACTTTCTGAGGCTGGGAAAGTATTTGGAGGCTTTTGTCTCCATGTC 6060
Db 6001 AAAATTTATGAAAACTTTCTGAGGCTGGGAAAGTATTTGGAGGCTTTTGTCTCCATGTC 6060
Qy 6061 CAGGTTCACTTACATCAATAAATTTCTTAAATGAGTATTTGCTTTCAATTAGCAACA 6120
Db 6061 CAGGTTCACTTACATCAATAAATTTCTTAAATGAGTATTTGCTTTCAATTAGCAACA 6120
Qy 6121 TATGCTTCAAGGAAAGGACATAGTCAATCTGTTTATGCTAGTATTTTCCAGGA 6180
Db 6121 TATGCTTCAAGGAAAGGACATAGTCAATCTGTTTATGCTAGTATTTTCCAGGA 6180
Qy 6181 ATTTATTTCCCTCTCAATTTTGTCTCAATTTTCAATTTTATTTTCCATCCATGAGTGAAG 6240
Db 6181 ATTTATTTCCCTCTCAATTTTGTCTCAATTTTCAATTTTATTTTCCATCCATGAGTGAAG 6240
Qy 6241 TCAGGTCAGAGTTGTAGACATTTTATGTTGTGTTAACTCTTCTGCAATTTTGTATTT 6300
Db 6241 TCAGGTCAGAGTTGTAGACATTTTATGTTGTGTTAACTCTTCTGCAATTTTGTATTT 6300
Qy 6301 GGTGTTTT 6308
Db 6301 GGTGTTTT 6308

RESULT 3
US-10-287-218-22
; Sequence 22, Application US/10287218
; Publication NO. US20030198975A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: AZIMZAI, Yalda; AU-YOUNG, Janice K.
; APPLICANT: BATRA, Sajeesh; BAUGHN, Mariah R.
; APPLICANT: BECHA, Shanya D.; BOROWSKI, Mark L.
; APPLICANT: BUFORD, Neil; DING, Li
; APPLICANT: ELLIOTT, Vicki S.; EMERLING, Brooke M.
; APPLICANT: GANDHI, Ameena R.; GIETZEN, Kimberly J.
; APPLICANT: GRIFFIN, Jennifer A.; HAFALIA, April J.A.
; APPLICANT: HONCHELL, Cynthia D.; LAL, Preti G.
; APPLICANT: LEE, Soo Yeun; LU, Dyung Aina M.
; APPLICANT: ARVIZU, Chandra S.; RAMKUMAR, Jayalaxmi
; APPLICANT: REDDY, Roopa; SANJANWALA, Madhu, M.
; APPLICANT: TANG, Y. Tom; WALIA, Narinder K.
; APPLICANT: WANG, Yu-mei, E.; WARREN, Bridget A.
; APPLICANT: XU, Yuming; YANG, Junming
; APPLICANT: YAO, Monique G.; YUE, Henry
; APPLICANT: ZEBARJADIAN, Yeganeh
; TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH
; FILE REFERENCE: PI-0417 USA
; CURRENT APPLICATION NUMBER: US/10/287,218
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: PCT/US02/11152
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: US 60/349,705
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: US 60/295,263
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/295,340
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/293,727
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/291,846
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/291,662
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: US 60/287,228
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/286,820
; PRIOR FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: US 60/283,294
; PRIOR FILING DATE: 2001-04-11
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 22
; LENGTH: 7742
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030198975A1 1351608CBI
US-10-287-218-22
Query Match 88.6%; Score 5592; DB 16; Length 7742;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5592; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 717 GGTACTGAGAGATGGAATCAGCGGAGTTACCCAGACCCCTCAGCGTCTGGCATCT 776
Db 1 GGTACTGAGAGATGGAATCAGCGGAGTTACCCAGACCCCTCAGCGTCTGGCATCT 60
QY 777 TGGTGGGATCAGCAAGTTGATTTTATCTGCTTTCTGTCATCTTGGCAAAATTTGGT 836
Db 61 TGGTGGGATCAGCAAGTTGATTTTATCTGCTTTCTGTCATCTTGGCAAAATTTGGT 120
QY 837 CCAGAAATTTACTTTGCTGAAATGGACCCAGACTTGGAAAGCAGGAGAAAGTGTACAA 896

Db 121 CCAGAAATTTACTTTGCTGAAATGGACCCAGACTTGGAAAGCAGGAGAAAGTGTACAA 180
QY 897 ATGTCAATATTTCACTCCACTGGATGTACTTATTTGGAGAGATCCAGATATTTGCTTA 956
Db 181 ATGTCAATATTTCACTCCACTGGATGTACTTATTTGGAGAGATCCAGATATTTGCTTA 240
QY 957 GAGAAATTTGAAGCACAGTGGAGCAATTTTCAGCTTTGTGGAGGGTTCCTTCAAAAGTGGAG 1016
Db 241 GAGAAATTTGAAGCACAGTGGAGCAATTTTCAGCTTTGTGGAGGGTTCCTTCAAAAGTGGAG 300
QY 1017 ACAACCTATTTCTGACGGATTTGCAATGTATCCAAATGTATCTGCTGATGAGTGC 1076
Db 301 ACAACCTATTTCTGACGGATTTGCAATGTATCCAAATGTATCTGCTGATGAGTGC 360
QY 1077 TTCCAGGACAGTGTTCATTAATAATCATCGTTTACAGATGCATCTTCTTCTGAGGAGGG 1136
Db 361 TTCCAGGACAGTGTTCATTAATAATCATCGTTTACAGATGCATCTTCTTCTGAGGAGGG 420
QY 1137 TTCTGTGACTGTGGAGACACAGAGGCATGGAAACTGGCCCTTTTGTGTAAATCATGAA 1196
Db 421 TTCTGTGACTGTGGAGACACAGAGGCATGGAAACTGGCCCTTTTGTGTAAATCATGAA 480
QY 1197 CCTGGAAGACAGTACTATATAAAGAGAAATTCACGCTGTCCGTTGGAATGAAGAGTAA 1256
Db 481 CCTGGAAGACAGTACTATATAAAGAGAAATTCACGCTGTCCGTTGGAATGAAGAGTAA 540
QY 1257 GTCCAGCCAGGAAATATTTCTTCAGTGATATAAATATGTCGTAGAAATCATCTATATGG 1316
Db 541 GTCCAGCCAGGAAATATTTCTTCAGTGATATAAATATGTCGTAGAAATCATCTATATGG 600
QY 1317 GAAGAGGAAAGAAACTGCTCCCTGAACTCCAGATAAGGGAGAAATAAGAAAGATCATAT 1376
Db 601 GAAGAGGAAAGAAACTGCTCCCTGAACTCCAGATAAGGGAGAAATAAGAAAGATCATAT 660
QY 1377 TGTGTCTCTTTTCAATGATGAACACCAATTCATATGACACCGTCATATACAGCCTACAA 1436
Db 661 TGTGTCTCTTTTCAATGATGAACACCAATTCATATGACACCGTCATATACAGCCTACAA 720
QY 1437 GCTCTTGACTGTGAGCTGCGAGAGCCCGAGTTGATACCACTGCCATTCGACAAAGAGGGT 1496
Db 721 GCTCTTGACTGTGAGCTGCGAGAGCCCGAGTTGATACCACTGCCATTCGACAAAGAGGGT 780
QY 1497 CGTGGGCTGTAAAAGCGGAGCTTATGCTCTTGCAGAGGAGCAAGAAAGAGATATAAG 1556
Db 781 CGTGGGCTGTAAAAGCGGAGCTTATGCTCTTGCAGAGGAGCAAGAAAGAGATATAAG 840
QY 1557 AGTCATTCAGAAATGTCTCTCAACATCCACTTCATGTAGAAGTATTACATCAGAGATT 1616
Db 841 AGTCATTCAGAAATGTCTCTCAACATCCACTTCATGTAGAAGTATTACATCAGAGATT 900
QY 1617 ATGGCTCATCAGAAATTTGCTTTGGCTTGGTCTTGGATGAAACAAATTTATGAGCTAT 1676
Db 901 ATGGCTCATCAGAAATTTGCTTTGGCTTGGTCTTGGATGAAACAAATTTATGAGCTAT 960
QY 1677 TCAAGTGACTTTAGCGAGATCTTTTGGCAAGATGCTTTAGAGAGAAACCTGACTCGGAG 1736
Db 961 TCAAGTGACTTTAGCGAGATCTTTTGGCAAGATGCTTTAGAGAGAAACCTGACTCGGAG 1020
QY 1737 AATCCCTGTCTCATAGCAGGTTAATGCTTTGGGATGCAAGCTTTTAAAGGTGCCCGT 1796
Db 1021 AATCCCTGTCTCATAGCAGGTTAATGCTTTGGGATGCAAGCTTTTAAAGGTGCCCGT 1080
QY 1797 AAGATCTCTCATGAATTTGATCTTCAGCAGTTTTTTATGGAGATGGATACAAAAATCTC 1856
Db 1081 AAGATCTCTCATGAATTTGATCTTCAGCAGTTTTTTATGGAGATGGATACAAAAATCTC 1140
QY 1857 TTTGCTATGGAATTTTGTGAAGTATTATAAACAACCTGCAGAAAGATATATCATGATGAT 1916
Db 1141 TTTGCTATGGAATTTTGTGAAGTATTATAAACAACCTGCAGAAAGATATATCATGATGAT 1200
QY 1917 CATCAGAGAGATGTCTCTATAAATCTGCACTTTTCAGTTCAGATGTTTACCTGTTCTCT 1976

Db 1201 CATGACAGAGTATCTCTATAACTGCACCTTTTCAGTTCAGATGTTTACTGTTCTCTACTCTG 1260
Qy 1977 GCTCGACATCTTATTGAAGACAGAAATGTTATCTCTGTCATTTACTGAAATCTCTGCTAGAA 2036
Db 1261 GCTCGACATCTTATTGAAGACAGAAATGTTATCTCTGTCATTTACTGAAATCTCTGCTAGAA 1320
Qy 2037 GTTTTACCTGAGTACTTGGACAGGAAACAAATAATTCACCTCCAGGTTTATAGCCAGGAC 2096
Db 1321 GTTTTACCTGAGTACTTGGACAGGAAACAAATAATTCACCTCCAGGTTTATAGCCAGGAC 1380
Qy 2097 AAATGGGAAGAGTATATGTCAGTAAATATGTCAGTAAAGTATATCTGATCAGCAAAACC 2156
Db 1381 AAATGGGAAGAGTATATGTCAGTAAATATGTCAGTAAAGTATATCTGATCAGCAAAACC 1440
Qy 2157 ACATATGACAGAAAGATTAAAGATGCGATTCCTTGAAGTTTTCGATCTTTTGAAG 2216
Db 1441 ACATATGACAGAAAGATTAAAGATGCGATTCCTTGAAGTTTTCGATCTTTTGAAG 1500
Qy 2217 ATTCTTACTGTATGACGGAATGGAAGAAATCCGAAGACAGGTTGGGCAACACATTGAA 2276
Db 1501 ATTCTTACTGTATGACGGAATGGAAGAAATCCGAAGACAGGTTGGGCAACACATTGAA 1560
Qy 2277 GTGGATCCTGATTGGGAGGCTGCCATTGCTATACAGATGCAATTCGAAGAAATATTTTACTC 2336
Db 1561 GTGGATCCTGATTGGGAGGCTGCCATTGCTATACAGATGCAATTCGAAGAAATATTTTACTC 1620
Qy 2337 ATGTTCCAGAGTGTGTGTCATGGAAGAACTCTTACTTGGCTTTAAGAAATGT 2396
Db 1621 ATGTTCCAGAGTGTGTGTCATGGAAGAACTCTTACTTGGCTTTAAGAAATGT 1680
Qy 2397 CACAAGCTGTGATGAGTGCAGTACAGTTCATATCTAGTAGCAAGACAGTAGTACAA 2456
Db 1681 CACAAGCTGTGATGAGTGCAGTACAGTTCATATCTAGTAGCAAGACAGTAGTACAA 1740
Qy 2457 TCGTGTGACATAGTTTGGAAACAAAGTCTACAGAGTATCTGAGATCTTGTGAAGCATA 2516
Db 1741 TCGTGTGACATAGTTTGGAAACAAAGTCTACAGAGTATCTGAGATCTTGTGAAGCATA 1800
Qy 2517 CATCTGCCATCTCTAGAACCTTCTGCTGCTTCATGTAAGTAAAGCAGCTGGGTGCT 2576
Db 1801 CATCTGCCATCTCTAGAACCTTCTGCTGCTTCATGTAAGTAAAGCAGCTGGGTGCT 1860
Qy 2577 GTTTTCAAGCTGCATGAAATTTGTGTCTTTTGGAGCTTTTCAAGTAGAGGTACTAGTGAA 2636
Db 1861 GTTTTCAAGCTGCATGAAATTTGTGTCTTTTGGAGCTTTTCAAGTAGAGGTACTAGTGAA 1920
Qy 2637 TATCTTTTACGTTGTCTGTTGTGTCGCCAGGTTGTTGCTGAGATGTGGCGAAGAAAT 2696
Db 1921 TATCTTTTACGTTGTCTGTTGTGTCGCCAGGTTGTTGCTGAGATGTGGCGAAGAAAT 1980
Qy 2697 GGACTGTCTCTTATTAGCCAGGTGTTTATTACCAAGATGTTTAAAGTCAGAGAGAAATG 2756
Db 1981 GGACTGTCTCTTATTAGCCAGGTGTTTATTACCAAGATGTTTAAAGTCAGAGAGAAATG 2040
Qy 2757 TATGATAAAGATATCATATGCTTCAGATTGGTGCATCTTTAAAGTATGCCATTAAGTTC 2816
Db 2041 TATGATAAAGATATCATATGCTTCAGATTGGTGCATCTTTAAAGTATGCCATTAAGTTC 2100
Qy 2817 TTGTTTACTGTTTCTCAGAGGTATGAACTTGGCCAGGCTTTTAAAGACCATATCTACA 2876
Db 2101 TTGTTTACTGTTTCTCAGAGGTATGAACTTGGCCAGGCTTTTAAAGACCATATCTACA 2160
Qy 2877 AAAGACCAAGGATTTGATTAACATATATATACATTAATAGAGAAATGCTTCAGGTCTC 2936
Db 2161 AAAGACCAAGGATTTGATTAACATATATATACATTAATAGAGAAATGCTTCAGGTCTC 2220
Qy 2937 ATCTATATGTTGGGTGAGGTTATGTAATCTGAGTGGGAATGTAACAAAGAGAGGTC 2996
Db 2221 ATCTATATGTTGGGTGAGGTTATGTAATCTGAGTGGGAATGTAACAAAGAGAGGTC 2280
Qy 2997 ACNATGAGAAATCATTTCACTTGTCTTGGATTGAAACCCATGCCACAGTGCCATTTGCC 3056
Db 2281 ACNATGAGAAATCATTTCACTTGTCTTGGATTGAAACCCATGCCACAGTGCCATTTGCC 2340

Qy 3057 AAAAATTTTACCTGAGAAATGAATAATGAATACTGCTTTAGAGATGTCATAAACAAAGTG 3116
Db 2341 AAAAATTTTACCTGAGAAATGAATAATGAATACTGCTTTAGAGATGTCATAAACAAAGTG 2400
Qy 3117 GCCACATTTTAAAGAAACCAAGGTGTATCAGGCCATGAGTTTATGAACCTAAAGAGTAATCA 3176
Db 2401 GCCACATTTTAAAGAAACCAAGGTGTATCAGGCCATGAGTTTATGAACCTAAAGAGTAATCA 2460
Qy 3177 CTGAAGAATCTTCAATATGTAATCTTTTATCACTCAAAACCAGCATAGCAAGGCTGAA 3236
Db 2461 CTGAAGAATCTTCAATATGTAATCTTTTATCACTCAAAACCAGCATAGCAAGGCTGAA 2520
Qy 3237 CATATGACAGAAAGAAAGGAGAAACCAAGAAACCAAGATGAAGCATTCGCCGCCACCA 3296
Db 2521 CATATGACAGAAAGAAAGGAGAAACCAAGAAACCAAGATGAAGCATTCGCCGCCACCA 2580
Qy 3297 CCTCTGAATTTCTGCCCTGCTTTTCAGCAAGAGTATTAACCTTCTCAACTGTGATATCATG 3356
Db 2581 CCTCTGAATTTCTGCCCTGCTTTTCAGCAAGAGTATTAACCTTCTCAACTGTGATATCATG 2640
Qy 3357 ATGTACATTTCTCAGGACCGTATTTTGAAGCGGCAATAGACACAGATTTCTAACTTTGAGACC 3416
Db 2641 ATGTACATTTCTCAGGACCGTATTTTGAAGCGGCAATAGACACAGATTTCTAACTTTGAGACC 2700
Qy 3417 GAAGGATGCTCCAAATGCGCTTTTCATATCTGGCATTTGGGTTTACTAGAGAGAGCA 3476
Db 2701 GAAGGATGCTCCAAATGCGCTTTTCATATCTGGCATTTGGGTTTACTAGAGAGAGCA 2760
Qy 3477 CAGCTTCAAAAGCTCTTGAAGAGAGTAACTTTGACTTTTATCATAGGCTTCAAGA 3536
Db 2761 CAGCTTCAAAAGCTCTTGAAGAGAGTAACTTTGACTTTTATCATAGGCTTCAAGA 2820
Qy 3537 TTGGAGGTTTCAAGCATGAAATATACAAATGCTTTTGGAAAAAAGCTCAAGAGAAATTCGCCAG 3596
Db 2821 TTGGAGGTTTCAAGCATGAAATATACAAATGCTTTTGGAAAAAAGCTCAAGAGAAATTCGCCAG 2880
Qy 3597 TTGAAGGCCAGAGAGGATGATTAACGTTGATATCTCAGATGTTTGAACAGTGAAGCA 3656
Db 2881 TTGAAGGCCAGAGAGGATGATTAACGTTGATATCTCAGATGTTTGAACAGTGAAGCA 2940
Qy 3657 TTGAAGAGAAATCTTGTAAATTTGTAGCAACCATCAGGATCGGAATCTATTAAAGAT 3716
Db 2941 TTGAAGAGAAATCTTGTAAATTTGTAGCAACCATCAGGATCGGAATCTATTAAAGAT 3000
Qy 3717 GATCAGATTTACTCATGATAGAAAAAGACAGAAACGAAAAAGAAAGCTGAAAGCTGCTAGG 3776
Db 3001 GATCAGATTTACTCATGATAGAAAAAGACAGAAACGAAAAAGAAAGCTGAAAGCTGCTAGG 3060
Qy 3777 CTACATGCCAGAGATCATGGCTCAGATGCTGCTTTACAGAAAAAAGCTTCAATTGAAACT 3836
Db 3061 CTACATGCCAGAGATCATGGCTCAGATGCTGCTTTACAGAAAAAAGCTTCAATTGAAACT 3120
Qy 3837 CATAACTCATGATGATGATACATCAGAAATGCTGGGAAGAGATTCATTTATGAG 3896
Db 3121 CATAACTCATGATGATGATACATCAGAAATGCTGGGAAGAGATTCATTTATGAG 3180
Qy 3897 GAAGAGACACCCAGCAGCTCAGTGAATCTTGAAGATTTGCTTTGGGCTTAAAGCGGCT 3956
Db 3181 GAAGAGACACCCAGCAGCTCAGTGAATCTTGAAGATTTGCTTTGGGCTTAAAGCGGCT 3240
Qy 3957 CCATCTGTTACTGAAAAAGGAGGTCCTGACGCTGCTTTCAGAGAAATGCTTCCAGAGAGAA 4016
Db 3241 CCATCTGTTACTGAAAAAGGAGGTCCTGACGCTGCTTTCAGAGAAATGCTTCCAGAGAGAA 3300
Qy 4017 AAAATAGAAATTAATGCTGATTTATCGGCTGTGTCAGAAATCTACTGCTTAAACC 4076
Db 3301 AAAATAGAAATTAATGCTGATTTATCGGCTGTGTCAGAAATCTACTGCTTAAACC 3360
Qy 4077 CAGCACAGGGGAAACCCATAGAACTCTCAGAGAGAGCCCTAGACCCACTTTTTCATGAT 4136
Db 3361 CAGCACAGGGGAAACCCATAGAACTCTCAGAGAGAGCCCTAGACCCACTTTTTCATGAT 3420

QY 4137 CCAGACTGGCATATGGAACCTTATACAGGAAGCTGTGGTCATGTAAATGCAAGCAGTGTGC 4196
DB |||||
DB 3421 CCAGACTTGGCATATGGAACCTTATACAGGAAGCTGTGGTCATGTAAATGCAAGCAGTGTGC 3480
QY 4197 TGGCAGAGATATTTTGAAGCTGTGACAGCTGAGCTCTCAGAGCGGCAATCATGTGTGACCTT 4256
DB |||||
DB 3481 TGGCAGAGATATTTTGAAGCTGTGACAGCTGAGCTCTCAGAGCGGCAATCATGTGTGACCTT 3540
QY 4257 TTTGACTTGGAAAGTGGAGATATCTTTGGCCCTCTTTGCAATCTCTGTGCAATACTGTG 4316
DB |||||
DB 3541 TTTGACTTGGAAAGTGGAGATATCTTTGGCCCTCTTTGCAATCTCTGTGCAATACTGTG 3600
QY 4317 ATCCCCATTAATCTTTGCAACCTCAAAAGATATAACAGTGAAGTGAAGTGTCTCTGTCT 4376
DB |||||
DB 3601 ATCCCCATTAATCTTTGCAACCTCAAAAGATATAACAGTGAAGTGAAGTGTCTCTGTCT 3660
QY 4377 CAACTTTTGAACCTTGCCACGGTGGATACAGACTGTCTTGGCCAGAAATATCAGGTTATAAT 4436
DB |||||
DB 3661 CAACTTTTGAACCTTGCCACGGTGGATACAGACTGTCTTGGCCAGAAATATCAGGTTATAAT 3720
QY 4437 ATAAGACATGCTAAAGGAGAAACCAATTCCTATTTTCTTTAATCAAGGAATGGAGAT 4496
DB |||||
DB 3721 ATAAGACATGCTAAAGGAGAAACCAATTCCTATTTTCTTTAATCAAGGAATGGAGAT 3780
QY 4497 TCTACTTTGGAGTTCATTTCCATCTCGAGTTTGGCGTTGAGTCTTCGATTAAATATTCA 4556
DB |||||
DB 3781 TCTACTTTGGAGTTCATTTCCATCTCGAGTTTGGCGTTGAGTCTTCGATTAAATATTCA 3840
QY 4557 AATAGCATCAAGGAAATGGTATTCTTTTGGCCACAACAATTTATAGAAATGGATGAAA 4616
DB |||||
DB 3841 AATAGCATCAAGGAAATGGTATTCTTTTGGCCACAACAATTTATAGAAATGGATGAAA 3900
QY 4617 GTGCCACCTGATGAAGGATCTCGAGTCCCATGCTGACCTGAGCAGCACCTGCGCTTTC 4676
DB |||||
DB 3901 GTGCCACCTGATGAAGGATCTCGAGTCCCATGCTGACCTGAGCAGCACCTGCGCTTTC 3960
QY 4677 ACTATCTCAGGCAATTTGAAATCTATTGGAGATGAAGGAAACCTCTGTTTGGAGCACTT 4736
DB |||||
DB 3961 ACTATCTCAGGCAATTTGAAATCTATTGGAGATGAAGGAAACCTCTGTTTGGAGCACTT 4020
QY 4737 CAAATATGACAGCATTAATGGTCTGAAGCATTAATGCAAGTGTGCAAGAGGAT 4796
DB |||||
DB 4021 CAAATATGACAGCATTAATGGTCTGAAGCATTAATGCAAGTGTGCAAGAGGAT 4080
QY 4797 ACCTGTCTCAGTCTGATACAGAAATCTGGTTCGTCTCTATCAGTGTCTTCTCT 4856
DB |||||
DB 4081 ACCTGTCTCAGTCTGATACAGAAACATCTGGTTCGTCTCTATCAGTGTCTTCTCT 4140
QY 4857 AACATAAATCAGAAGATACACATGCTCTGTCTATAGATCTGTTCATGTGTTGGTG 4916
DB |||||
DB 4141 AACATAAATCAGAAGATACACATGCTCTGTCTATAGATCTGTTCATGTGTTGGTG 4200
QY 4917 GGTGTGTGTAGCATTCCTCATCTGTATATGGAGTACCCCTGTGTGATCTGAGCCTTCT 4976
DB |||||
DB 4201 GGTGTGTGTAGCATTCCTCATCTGTATATGGAGTACCCCTGTGTGATCTGAGCCTTCT 4260
QY 4977 TCAGTGTGTCTTCTATACACCTTTATCTCTTCCATTTGATCACCATGSCACACATG 5036
DB |||||
DB 4261 TCAGTGTGTCTTCTATACACCTTTATCTCTTCCATTTGATCACCATGSCACACATG 4320
QY 5037 CTTTCAGATACTTACAGTAGACACAGGCTTACCCCTTGTCTCAGGTTCAAGAAGACAGT 5096
DB |||||
DB 4321 CTTTCAGATACTTACAGTAGACACAGGCTTACCCCTTGTCTCAGGTTCAAGAAGACAGT 4380
QY 5097 GAAGAGGCTCAATTCGCGCATCTTTCTTTTGCAGAAATTTCTCAATATACAGTGGCTCC 5156
DB |||||
DB 4381 GAAGAGGCTCAATTCGCGCATCTTTCTTTTGCAGAAATTTCTCAATATACAGTGGCTCC 4440
QY 5157 ATTGGGTGTGATATTCCTGGCTGGTATTTGTGGGTCTCACTGAAATGGCATCACCCCT 5216
DB |||||
DB 4441 ATTGGGTGTGATATTCCTGGCTGGTATTTGTGGGTCTCACTGAAATGGCATCACCCCT 4500
QY 5217 TATCTGTGTGTGCTGATTTGTTTTTCCACTATTTTACTTGGGGTAACCTCCGCTCAGGAA 5276

DB 4501 TATCTGTGTGTGCTGCAATTTTTCACATTAATTTACTTGGGGTAACCTCCGCTCAGGAA 4560
QY 5277 CTGCATACCAATTTCTGCAGAGGAGAGTACAGTGCACCTCTGTAGCTATCTATCTTTTACCT 5336
DB 4561 CTGCATACCAATTTCTGCAGAGGAGAGTACAGTGCACCTCTGTAGCTATCTATCTTTTACCT 4620
QY 5337 ACAAATTTGTTCTGCTCTTCCAGGAATATTTGGGATACTGTAAAGGCCCTTGTCTCCAGAGG 5396
DB 4621 ACAAATTTGTTCTGCTCTTCCAGGAATATTTGGGATACTGTAAAGGCCCTTGTCTCCAGAGG 4680
QY 5397 TGGTGTGAGATCTGCTCTTAACTTAACTGTTTGAAGCAAAAAACAACGTTGGTCAAGTAC 5456
DB 4681 TGGTGTGAGATCTGCTCTTAACTTAACTGTTTGAAGCAAAAAACAACGTTGGTCAAGTAC 4740
QY 5457 CTTAGAAAAAAGAAATAGTTTGTATAGAGCTTCTGATGATATATAGTCTGCTCTGATCAAA 5516
DB 4741 CTTAGAAAAAAGAAATAGTTTGTATAGAGCTTCTGATGATATATAGTCTGCTCTGATCAAA 4800
QY 5517 GCTTCTCATTTTCAGGTGCCACGGTCTGCAGATGATGAGCGAAAAAGCATCTGTCTCTGTC 5576
DB 4801 GCTTCTCATTTTCAGGTGCCACGGTCTGCAGATGATGAGCGAAAAAGCATCTGTCTCTGTC 4860
QY 5577 CTTTCTGTGGGCTATATATATGTCTCAGAAACATTTGTCTGCCAGGAAATTTGTGAACGGG 5636
DB 4861 CTTTCTGTGGGCTATATATATGTCTCAGAAACATTTGTCTGCCAGGAAATTTGTGAACGGG 4920
QY 5637 GAAGAGGTGTGAGCTTGCATTTTTCAGCACTTCACTGTGGAGCGGAGTCTGCATTTTC 5696
DB 4921 GAAGAGGTGTGAGCTTGCATTTTTCAGCACTTCACTGTGGAGCGGAGTCTGCATTTTC 4980
QY 5697 CTAATAATCAGAGAAATGCCAGTGTCTGTGTTGAAGTAAAGCCAGAGGCTGTGCCTAT 5756
DB 4981 CTAATAATCAGAGAAATGCCAGTGTCTGTGTTGAAGTAAAGCCAGAGGCTGTGCCTAT 5040
QY 5757 CCAGCTCTTACTTTGGATGAATATGAGAAACAGACCTTGCCCTGGAAGGGGCAACCCC 5816
DB 5041 CCAGCTCTTACTTTGGATGAATATGAGAAACAGACCTTGCCCTGGAAGGGGCAACCCC 5100
QY 5817 CTTCAATTTACTCTGTGAGCGGTATCGGAGCTTCCATTTGCTCTGCAACACACTGCAT 5876
DB 5101 CTTCAATTTACTCTGTGAGCGGTATCGGAGCTTCCATTTGCTCTGCAACACACTGCAT 5160
QY 5877 ATAGAAGAGATTTGCTAGGAGCAAGAGACTAATCAGATGTTATTTGGATTCAACTGGCAG 5936
DB 5161 ATAGAAGAGATTTGCTAGGAGCAAGAGACTAATCAGATGTTATTTGGATTCAACTGGCAG 5220
QY 5937 TTACTGTGAGCTCCAACTCTGCTCAAGACAAATCAAAATGCAAGTGTAAAGGCTG 5996
DB 5221 TTACTGTGAGCTCCAACTCTGCTCAAGACAAATCAAAATGCAAGTGTAAAGGCTG 5280
QY 5997 ATTCAAAATTAAGAAATCTTCTGAGGCTGCGGAAAGTATTTGGGGCTCTTTTGTCTCA 6056
DB 5281 ATTCAAAATTAAGAAATCTTCTGAGGCTGCGGAAAGTATTTGGGGCTCTTTTGTCTCA 5340
QY 6057 TGTCCAGGCTCACTTACATCAATAAATAATTTCTTAATGAGTATTTGCTTCAATTAGCA 6116
DB 5341 TGTCCAGGCTCACTTACATCAATAAATAATTTCTTAATGAGTATTTGCTTCAATTAGCA 5400
QY 6117 AACATATGCTTTCAGGAAAAAAGACATAGATCAATCTGTTTTTATGTGTAGTATTTTCC 6176
DB 5401 AACATATGCTTTCAGGAAAAAAGACATAGATCAATCTGTTTTTATGTGTAGTATTTTCC 5460
QY 6177 AGGAATTTATTTCCCTTTCATTAATTTGTCTCAATTTTATTTTTCATCTTCCTGATAGT 6236
DB 5461 AGGAATTTATTTCCCTTTCATTAATTTGTCTCAATTTTATTTTTCATCTTCCTGATAGT 5520
QY 6237 GAAGTCACTCAAAACAGTTGTAGACATTTTATGTGTGTTAACTCTTCTGCAATTTTGT 6296
DB 5521 GAAGTCACTCAAAACAGTTGTAGACATTTTATGTGTGTTAACTCTTCTGCAATTTTGT 5580
QY 6297 ATTTGGTGTTTT 6308
DB |||||

Db 5581 ATTGGTGTGTTT 5592

RESULT 4
US-10-474-291-22
; Sequence 22, Application US/10474291
; Publication No. US20040132043A1
; GENERAL INFORMATION:
; APPLICANT: AZIMZAI, Yalda; AU-YOUNG, Janice K.
; APPLICANT: BATRA, Sajeev; BAUGHN, Mariah R.
; APPLICANT: BECHA, Shanya D.; BOROWSKY, Mark L.
; APPLICANT: BURFORD, Neil; DING, Li
; APPLICANT: ELLIOTT, Vicki S.; EMERLING, Brooke M.
; APPLICANT: GANDHI, Ameena R.; GIETZEN, Kimberly J.A.
; APPLICANT: GRIFFIN, Jennifer A.; HAPALIA, April J.A.
; APPLICANT: HONCHELL, Cynthia D.; LAL, Preeti G.
; APPLICANT: LEE, Soo Yeun; LU, Dyung Aina M.
; APPLICANT: ARVIZU, Chandra S.; RAMKUMAR, Jayalaxmi
; APPLICANT: REDDY, Roopa M.; SANTANWALA, Madhusudan M.
; APPLICANT: TANG, Y. Tom; CHAWLA, Narinder K.
; APPLICANT: WANG, Yu-Mei E.; WARREN, Bridget A.
; APPLICANT: XU, Yuming; YANG, Junming
; APPLICANT: YAO, Monique G.; YUE, Henry
; APPLICANT: ZEBARJADIAN, Yeganeh
; TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH
; FILE REFERENCE: PI-0417 USN
; CURRENT APPLICATION NUMBER: US/10/474,291
; CURRENT FILING DATE: 2003-10-06
; PRIOR APPLICATION NUMBER: PCT/US02/11152
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: US 60/349,705
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: US 60/295,263
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/295,340
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/293,727
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/291,846
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/291,662
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: US 60/287,228
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/286,820
; PRIOR FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: US 60/283,294
; PRIOR FILING DATE: 2001-04-11
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 22
; LENGTH: 7742
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 1351608CB1
US-10-474-291-22.

Query Match 88.6%; Score 5592; DB 19; Length 7742;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 5592; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 717 GGTACTGAGGATGGAATCAGCGGAGTTACCCAGACCCCTCAGCGTCTGCCATCT 776
Db 1 GGTACTGAGAGATGGAATCAGCGGAGTTACCCAGACCCCTCAGCGTCTGCCATCT 60

Qy 777 TGGTGGGATCAGCAAGTTGATTTTATCTGCTTTCTTGCAATCATTTGGCACAATTTGGTG 836
Db 61 TGGTGGGATCAGCAAGTTGATTTTATCTGCTTTCTTGCAATCATTTGGCACAATTTGGTG 120

Qy 837 CCAGAAATTTACTTTTGTCTGAAATGACCCAGACTTTGGAAAGACGAGGAAAGTGTACAA 896

Db 121 CCAGAAATTTACTTTTGTCTGAAATGACCCAGACTTTGGAAAGACGAGGAAAGTGTACAA 180

Qy 897 ATGTCAATATTACATCCACTCGAATGGTACTTTATTTGGAGAGATCCAGATATTTGCTTTA 956

Db 181 ATGTCAATATTACATCCACTCGAATGGTACTTTATTTGGAGAGATCCAGATATTTGCTTTA 240

Qy 957 GAGAAATTTGAAGCACAGTGGAGCATTTTCAGCTTTTGTGGAGGGTTTCAAAGTGGAGAG 1016

Db 241 GAGAAATTTGAAGCACAGTGGAGCATTTTCAGCTTTTGTGGAGGGTTTCAAAGTGGAGAG 300

Qy 1017 ACAACCTATTCTTGGCAGGGATTGTGCAATTTGATCAACATGTGTAATCTGTATGACTGC 1076

Db 301 ACAACCTATTCTTGGCAGGGATTGTGCAATTTGATCAACATGTGTAATCTGTATGACTGC 360

Qy 1077 TTCCAGGACAGTGTTCATAAAATCATCGTTTACAAGATGCACTACTTTCTACTGGAGGAG 1136

Db 361 TTCCAGGACAGTGTTCATAAAATCATCGTTTACAAGATGCACTACTTTCTACTGGAGGAG 420

Qy 1137 TTCTGTGACTGTGGAGACACAGAGGCATGGAAAACCTGGCCCTTTTGTGTGTAATCATGAA 1196

Db 421 TTCTGTGACTGTGGAGACACAGAGGCATGGAAAACCTGGCCCTTTTGTGTGTAATCATGAA 480

Qy 1197 CCTGGAAGCAGAGTACTATATAAGAGAAATTCAGCTGTCCGTTGCAATGAAGAGGTAATT 1256

Db 481 CCTGGAAGCAGAGTACTATATAAGAGAAATTCAGCTGTCCGTTGCAATGAAGAGGTAATT 540

Qy 1257 GTCCAAAGCCAGGAAAAATTTTCCTTTCAGTGATAAAATATGTCTGTAAGAAATGACTATATGG 1316

Db 541 GTCCAAAGCCAGGAAAAATTTTCCTTTCAGTGATAAAATATGTCTGTAAGAAATGACTATATGG 600

Qy 1317 GAAGAGGAAAAAGAACTGCTCTGAACTCCAGATTAAGGAGAAAAATGAAGAGATACAT 1376

Db 601 GAAGAGGAAAAAGAACTGCTCTGAACTCCAGATTAAGGAGAAAAATGAAGAGATACAT 660

Qy 1377 TGTGTCTCTTTTCAATGATGAACACACATTCATATGACACGCTCATATACAGCTACAAAGA 1436

Db 661 TGTGTCTCTTTTCAATGATGAACACACATTCATATGACACGCTCATATACAGCTACAAAGA 720

Qy 1437 GCTCTTGACTGTGAGCTCGAGAGGCCAGTTTGCATACACTGCCATTCGACAAAGAGGT 1496

Db 721 GCTCTTGACTGTGAGCTCGAGAGGCCAGTTTGCATACACTGCCATTCGACAAAGAGGT 780

Qy 1497 CGTGGGCTGTAAAGCGGAGCTTATGCTGCTGCCAGGAGCAAGAAAGATATAAG 1556

Db 781 CGTGGGCTGTAAAGCGGAGCTTATGCTGCTGCCAGGAGCAAGAAAGATATAAG 840

Qy 1557 AGTCATTTCAGAAATGCTCTCAACATCCACTTCATGTAGAGTATTTACACTCAGAGATT 1616

Db 841 AGTCATTTCAGAAATGCTCTCAACATCCACTTCATGTAGAGTATTTACACTCAGAGATT 900

Qy 1617 ATGGCTCATCAGAAATTTGCTTGGCTCTTGGTTCCTGGATGAACAAATTTATGAGCTAT 1676

Db 901 ATGGCTCATCAGAAATTTGCTTGGCTCTTGGTTCCTGGATGAACAAATTTATGAGCTAT 960

Qy 1677 TCAAAGTGAATTTAGGACAGATCTTTTGGCAAGCATGCTCTTAGAGAGAAACCTGACTCGAG 1736

Db 961 TCAAAGTGAATTTAGGACAGATCTTTTGGCAAGCATGCTCTTAGAGAGAAACCTGACTCGAG 1020

Qy 1737 AATCCCTGTCTCATAGCAGGTTAATGCTTTGGGATGCAAGCTTTATAAGGTGCCCCGT 1796

Db 1021 AATCCCTGTCTCATAGCAGGTTAATGCTTTGGGATGCAAGCTTTATAAGGTGCCCCGT 1080

Qy 1797 AAGATCCTTCATGAATTTGATCTTCAGCAGTGTCTTATGGAGATGGAATACAAAACTC 1856

Db 1081 AAGATCCTTCATGAATTTGATCTTCAGCAGTGTCTTATGGAGATGGAATACAAAACTC 1140

Qy 1857 TTTGCTATGGAATTTGTGAAGTATTTATAAACAACCTGCAGAAAGAAATATATCAGTGATGAT 1916

Db 1141 TTTGCTATGGAATTTGTGAAGTATTTATAAACAACCTGCAGAAAGAAATATATCAGTGATGAT 1200

Qy 1917 CATGACAGAGATCTCTATTAACCTGCACTTTTCAGTTTCAGATGTTTACTGTTCTCTACTCTG 1976

Db 1201 CATGACAGAGTATCTCTATAACTGCACCTTTTCAGTTTCAGATGTTTACTGTTCTTACTCTG 1260
Qy 1977 GCTCGACATCTTATTGAAGAGCAGAAATGTTATCTCTCTCATTTACTGAACTCTGCTAGAA 2036
Db 1261 GCTCGACATCTTATTGAAGAGCAGAAATGTTATCTCTCTCATTTACTGAACTCTGCTAGAA 1320
Qy 2037 GTTTTACCTGAGTACTTGGACAGGAAACAATAAATCAACTCTCCAGGGTTATAGCCAGGAC 2096
Db 1321 GTTTTACCTGAGTACTTGGACAGGAAACAATAAATCAACTCTCCAGGGTTATAGCCAGGAC 1380
Qy 2097 AAATGGGAAGATATATGCGATTAATATGTGACCTTAAAGTATATCTGATCAGCAAAACC 2156
Db 1381 AAATGGGAAGATATATGCGATTAATATGTGACCTTAAAGTATATCTGATCAGCAAAACC 1440
Qy 2157 ACAATATGGACAGAAAGATTAAAGATGAGTTCTTGAAGTTTTCGATCTTTTGAAG 2216
Db 1441 ACAATATGGACAGAAAGATTAAAGATGAGTTCTTGAAGTTTTCGATCTTTTGAAG 1500
Qy 2217 ATTCTTACCTGTATGCGGGAATGGAAGAAATCCGAAGACAGGTTGGGCAACACATTGAA 2276
Db 1501 ATTCTTACCTGTATGCGGGAATGGAAGAAATCCGAAGACAGGTTGGGCAACACATTGAA 1560
Qy 2277 GTGATCTCTGATTTGGAGGCTGCCATTTGCTATACAGATGCAATTCGAAGAATATTTTACTC 2336
Db 1561 GTGGATCTCTGATTTGGAGGCTGCCATTTGCTATACAGATGCAATTCGAAGAATATTTTACTC 1620
Qy 2337 ATGTTCCAAAGTGTGTGCTGTGATGAGAACTCTTACTTGTGGCTTATAAAGAAATGT 2396
Db 1621 ATGTTCCAAAGTGTGTGCTGTGATGAGAACTCTTACTTGTGGCTTATAAAGAAATGT 1680
Qy 2397 CACAAGCTGTATGAGGTGAGTACACAGTTTCATATCTAGTAGCAAGCAGTAGTACAA 2456
Db 1681 CACAAGCTGTATGAGGTGAGTACACAGTTTCATATCTAGTAGCAAGCAGTAGTACAA 1740
Qy 2457 TCGTGTGGACATAGTTTGGAAACAAAGTCTTACAGAGTATCTGAGGATCTTTGTAAAGCATA 2516
Db 1741 TCGTGTGGACATAGTTTGGAAACAAAGTCTTACAGAGTATCTGAGGATCTTTGTAAAGCATA 1800
Qy 2517 CATCTGCCATCTCTAGGACCTTCTGCTGTCTTCACTGATACGTTTAAAGCAGGCTGGGTCT 2576
Db 1801 CATCTGCCATCTCTAGGACCTTCTGCTGTCTTCACTGATACGTTTAAAGCAGGCTGGGTCT 1860
Qy 2577 GTTTCGAAGCTCATGAAATTTGTCTTTTGAAGGACTTTCAAGTAGAGGTACTAGTGGAA 2636
Db 1861 GTTTCGAAGCTCATGAAATTTGTCTTTTGAAGGACTTTCAAGTAGAGGTACTAGTGGAA 1920
Qy 2637 TATCCTTTTACGTTGTCTGTGTGTGTCGCCAGGTTGTGTGAGATGTGGCGAAGAAAT 2696
Db 1921 TATCCTTTTACGTTGTCTGTGTGTGTCGCCAGGTTGTGTGAGATGTGGCGAAGAAAT 1980
Qy 2697 GGACTGTCTTTATTAGCCAGGTGTTTATTACCAAGATGTTTAAAGTCGAGAGAAATG 2756
Db 1981 GGACTGTCTTTATTAGCCAGGTGTTTATTACCAAGATGTTTAAAGTCGAGAGAAATG 2040
Qy 2757 TATGATAAAGATATCATCATGCTTCAGATTGTGTGATCTTTAATGATCCCAATTAAGTTC 2816
Db 2041 TATGATAAAGATATCATCATGCTTCAGATTGTGTGATCTTTAATGATCCCAATTAAGTTC 2100
Qy 2817 TTGTTACTGTTACTTCAGAGGTATGAACTTGGCCGAGGCTTTTAAACAAGACCAATATCTACA 2876
Db 2101 TTGTTACTGTTACTTCAGAGGTATGAACTTGGCCGAGGCTTTTAAACAAGACCAATATCTACA 2160
Qy 2877 AAAGACCCAGGATTTGAATTAACAATATACTAATAAGAGAAATGCTTCAGGTCTCTC 2936
Db 2161 AAAGACCCAGGATTTGAATTAACAATATACTAATAAGAGAAATGCTTCAGGTCTCTC 2220
Qy 2937 ATCTATATTGTGGTGAGGTTATGTACCTGAGGTGGGAAATGTGACCAAGAGAGGTTC 2996
Db 2221 ATCTATATTGTGGTGAGGTTATGTACCTGAGGTGGGAAATGTGACCAAGAGAGGTTC 2280
Qy 2997 ACAATGAGAGAAATCATCTTGCCTTGTGATTTGAACCCCATGCCACACAGTCCCATGCC 3056
Db 2281 ACAATGAGAGAAATCATCTTGCCTTGTGATTTGAACCCCATGCCACACAGTCCCATGCC 2340

Qy 3057 AAAAATTTTACCTGAGAATGAAAATAATAAGAACTGGCTTTAGAGAATGTCTATAAACAAGTGT 3116
Db 2341 AAAAATTTTACCTGAGAATGAAAATAATAAGAACTGGCTTTAGAGAATGTCTATAAACAAGTGT 2400
Qy 3117 GCCACATTTTAAAGAAACAGAGTGTATCAGGGCAGATGTTTATGAACTTAAAGATGAATCA 3176
Db 2401 GCCACATTTTAAAGAAACAGAGTGTATCAGGGCAGATGTTTATGAACTTAAAGATGAATCA 2460
Qy 3177 CTGAAAGACTTCAATATGTTACTTTTATCATTTACTCCAAACCCAGCATAGCAAGGCTGAA 3236
Db 2461 CTGAAAGACTTCAATATGTTACTTTTATCATTTACTCCAAACCCAGCATAGCAAGGCTGAA 2520
Qy 3237 CATATGCAAGAAAGAGGAGAAAACAAGAAAACAAGATGAAGCAATTTGCCGCCACCA 3296
Db 2521 CATATGCAAGAAAGAGGAGAAAACAAGAAAACAAGATGAAGCAATTTGCCGCCACCA 2580
Qy 3297 CCTCCTGAATTTCTGCCCTGCTTTTTCAGCAAAAGTGAATTAACCTTCTCAACTGTGATATCATG 3356
Db 2581 CCTCCTGAATTTCTGCCCTGCTTTTTCAGCAAAAGTGAATTAACCTTCTCAACTGTGATATCATG 2640
Qy 3357 ATGTACATTTCTCAGGACCGTATTTTTCAGGGGCAATAGACACAGATTTCTAACTTTGGAC 3416
Db 2641 ATGTACATTTCTCAGGACCGTATTTTTCAGGGGCAATAGACACAGATTTCTAACTTTGGAC 2700
Qy 3417 GAAGGGATGCTCCAAATGGCTTTTTCATATTTCTGGCATTTGGGTTTTACTAGAAAGAGCA 3476
Db 2701 GAAGGGATGCTCCAAATGGCTTTTTCATATTTCTGGCATTTGGGTTTTACTAGAAAGAGCA 2760
Qy 3477 CAGCTTCAAAAGCTCTCAGAAAGAAAGTAACTTTTGAATTTTATCAATAAGGCTTCAAGA 3536
Db 2761 CAGCTTCAAAAGCTCTCAGAAAGAAAGTAACTTTTGAATTTTATCAATAAGGCTTCAAGA 2820
Qy 3537 TTGGGAAGTTTCAGGCATGAAATATACAAATGCTTTTGGAAAACCTCAAAGGAATTTCCCGAG 3596
Db 2821 TTGGGAAGTTTCAGGCATGAAATATACAAATGCTTTTGGAAAACCTCAAAGGAATTTCCCGAG 2880
Qy 3597 TTAGAAGCCAGAGAGGACATGATAACGTTGGATATCTTCAGATGTTTTCACACAGTGAAGCGA 3656
Db 2881 TTAGAAGCCAGAGAGGACATGATAACGTTGGATATCTTCAGATGTTTTCACACAGTGAAGCGA 2940
Qy 3657 TTAAGAGAAAATCTTGTTTTAAATTTGTAGCAACCAATCAGGATCGGAATCTATTAAAGAA 3716
Db 2941 TTAAGAGAAAATCTTGTTTTAAATTTGTAGCAACCAATCAGGATCGGAATCTATTAAAGAA 3000
Qy 3717 GATGAGATTACTCATGATAAAGAAAGCAGAAACGAAAAGAAAGCTGAAGCTGTAGG 3776
Db 3001 GATGAGATTACTCATGATAAAGAAAGCAGAAACGAAAAGAAAGCTGAAGCTGTAGG 3060
Qy 3777 CTACATGCCAGAAAGATCATGGCTCAGATGTCTGCTTTACAGAAAACCTTCATTGAAAAT 3836
Db 3061 CTACATGCCAGAAAGATCATGGCTCAGATGTCTGCTTTACAGAAAACCTTCATTGAAAAT 3120
Qy 3837 CATAACTCATGTATGACAAATCATACAGAAATGCTCGGAAAGAAAGATTCATTATGAG 3896
Db 3121 CATAACTCATGTATGACAAATCATACAGAAATGCTCGGAAAGAAAGATTCATTATGAG 3180
Qy 3897 GAAGAGACCCCGAGGAGTCTACATGATGCTTCTAGAAATTTGGGTCTTAAACGGGGT 3956
Db 3181 GAAGAGACCCCGAGGAGTCTACATGATGCTTCTAGAAATTTGGGTCTTAAACGGGGT 3240
Qy 3957 CCATCTGTCTGAAAAGGAGGTCTGACGTGCTTTCGCAAGAAAGAAAGAGAGGTG 4016
Db 3241 CCATCTGTCTGAAAAGGAGGTCTGACGTGCTTTCGCAAGAAAGAAAGAGAGGTG 3300
Qy 4017 AAAATAGAAAATAATGCCATGTTATTCGGCTGTGTCCAGAAATCTATGCTCTTAACC 4076
Db 3301 AAAATAGAAAATAATGCCATGTTATTCGGCTGTGTCCAGAAATCTATGCTCTTAACC 3360
Qy 4077 CAGCACGGGAAAACCCCATAGAACTCTCAGAGAGAGCCCTAGACCCACTTTTCATGAT 4136
Db 3361 CAGCACGGGAAAACCCCATAGAACTCTCAGAGAGAGCCCTAGACCCACTTTTCATGAT 3420

Db	5581	ATTGGTGT	5592	
RESULT 5				
US-10-357-819-9				
; Sequence 9, Application US/10357819				
; Publication No. US20040259774A1				
; GENERAL INFORMATION:				
; APPLICANT: Alvarez, Enrique				
; APPLICANT: Edinger, Shlomit R.				
; APPLICANT: Gangolli, Esha A.				
; APPLICANT: Gerlach, Valerie L.				
; APPLICANT: Gorman, Linda				
; APPLICANT: Guo, Xiaojia				
; APPLICANT: Ji, Weizhen				
; APPLICANT: Kekuda, Ramesh				
; APPLICANT: Li, Li				
; APPLICANT: Miller, Charles E.				
; APPLICANT: Padigaru, Muralidhara				
; APPLICANT: Pattursajan, Meera				
; APPLICANT: Rastelli, Inca				
; APPLICANT: Rieger, Daniel K.				
; APPLICANT: Shenoy, Suresh G.				
; APPLICANT: Shimkets, Richard A.				
; APPLICANT: Spytek, Kimberly A.				
; APPLICANT: Zhong, Mei				
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD				
; FILE REFERENCE: 21402-538A				
; CURRENT APPLICATION NUMBER: US/10/357,819				
; PRIOR FILING DATE: 2003-02-03				
; PRIOR APPLICATION NUMBER: 09/520,781				
; PRIOR FILING DATE: 2000-03-08				
; PRIOR APPLICATION NUMBER: 09/584,411				
; PRIOR FILING DATE: 2000-05-31				
; PRIOR APPLICATION NUMBER: 09/783,436				
; PRIOR FILING DATE: 2001-02-14				
; PRIOR APPLICATION NUMBER: 10/085,198				
; PRIOR FILING DATE: 2002-02-25				
; PRIOR APPLICATION NUMBER: 60/353,301				
; PRIOR FILING DATE: 2002-02-01				
; PRIOR APPLICATION NUMBER: 60/355,099				
; PRIOR FILING DATE: 2002-02-08				
; PRIOR APPLICATION NUMBER: 60/356,424				
; PRIOR FILING DATE: 2002-02-12				
; PRIOR APPLICATION NUMBER: 60/358,239				
; PRIOR FILING DATE: 2002-02-20				
; PRIOR APPLICATION NUMBER: 60/358,608				
; PRIOR FILING DATE: 2002-02-21				
; PRIOR APPLICATION NUMBER: 60/359,367				
; PRIOR FILING DATE: 2002-02-25				
; Remaining Prior Application data removed - See File Wrapper or PALM.				
; NUMBER OF SEQ ID NOS: 142				
; SOFTWARE: CuraSeqList version 0.1				
; SEQ ID NO 9				
; LENGTH: 5466				
; TYPE: DNA				
; ORGANISM: Homo sapiens				
; FEATURE:				
; NAME/KEY: CDS				
; LOCATION: (66)..(5312)				
US-10-357-819-9				
Query Match 86.7%; Score 5466; DB 20; Length 5466;				
Best Local Similarity 100.0%; Pred. No. 0;				
Matches 5466; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Qy	631	CGGCCACAGGTTTCGGCTCTCGCGCGGGTCCGCACTCGCAGGCGTCAGTTTCCC	690	
Db	1	CGGCCACAGGTTTCGGCTCTCGCGCGGGTCCGCACTCGCAGGCGTCAGTTTCCC	60	
Qy	691	TCAAGATGCGGACGAGGAGGCTGGAGGTACTGAGAGGATGGAATCAGCGCGAGTTAC	750	
Db	61	TCAAGATGCGGACGAGGAGGCTGGAGGTACTGAGAGGATGGAATCAGCGCGAGTTAC	120	
Qy	751	CCGAGACCCCTCAGCGCTCGCATCTTGTGGGATCAGCAAGTTGATTTTTTATCTGCTT	810	
Db	121	CCGAGACCCCTCAGCGCTCGCATCTTGTGGGATCAGCAAGTTGATTTTTTATCTGCTT	180	
Qy	811	TCTTGCAATCAATTTGGCACAATTTGGTGCAGAAAATTTACTTTTGTGAAATGGACCCAGACT	870	
Db	181	TCTTGCAATCAATTTGGCACAATTTGGTGCAGAAAATTTACTTTTGTGAAATGGACCCAGACT	240	
Qy	871	TGAAAAAGCAGGAGGAAGTGTACAATGTCAATATTTCATCTCCAATGGAATGGTACTTAT	930	
Db	241	TGAAAAAGCAGGAGGAAGTGTACAATGTCAATATTTCATCTCCAATGGAATGGTACTTAT	300	
Qy	931	TTGGAGAGATCCAGATATTCTTGTAGAGAAATTTGAAGCAGCAGTGGAGCAATTCAGCTTT	990	
Db	301	TTGGAGAGATCCAGATATTCTTGTAGAGAAATTTGAAGCAGCAGTGGAGCAATTCAGCTTT	360	
Qy	991	GTGGAGGGTTTTTCAAAAGTGGAGAGACAACCTATTCTTGCAGGGATTTGTGCAATTTGATC	1050	
Db	361	GTGGAGGGTTTTTCAAAAGTGGAGAGACAACCTATTCTTGCAGGGATTTGTGCAATTTGATC	420	
Qy	1051	CAACATGTACTCTGTATGGACTCTTCCAGGACAGTGTTCATAAAAATCATCGTTACA	1110	
Db	421	CAACATGTACTCTGTATGGACTCTTCCAGGACAGTGTTCATAAAAATCATCGTTACA	480	
Qy	1111	AGATGCATCTTCTACTGGAGGAGGTTCTGTGACTGTGGAGACACAGAGCATGGAAAA	1170	
Db	481	AGATGCATCTTCTACTGGAGGAGGTTCTGTGACTGTGGAGACACAGAGCATGGAAAA	540	
Qy	1171	CTGGCCCTTTTGTGTAAATCATGAACCTGGAAGCAGGAGTACTATAAAGAGAAATTCAC	1230	
Db	541	CTGGCCCTTTTGTGTAAATCATGAACCTGGAAGCAGGAGTACTATAAAGAGAAATTCAC	600	
Qy	1231	GCTGTCCGTTGAATGGAAGAGTAATTTGTCCAAAGCAGGAAATATTTCTTCAGTGATAA	1290	
Db	601	GCTGTCCGTTGAATGGAAGAGTAATTTGTCCAAAGCAGGAAATATTTCTTCAGTGATAA	660	
Qy	1291	AATATGTCTGAGAAATGCACTATATGGGAAGGAAAAAGAACTGCTCTCAATGATGAACACCATCATATG	1350	
Db	661	AATATGTCTGAGAAATGCACTATATGGGAAGGAAAAAGAACTGCTCTCAATGATGAACACCATCATATG	720	
Qy	1351	TAAGGAGAAAAATGAAAGATACCTATTGTGTCTTTTCAATGATGAACACCATCATATG	1410	
Db	721	TAAGGAGAAAAATGAAAGATACCTATTGTGTCTTTTCAATGATGAACACCATCATATG	780	
Qy	1411	ACCAAGTCATATACAGCTTCAAAAGAGCTCTTGAGCTGTGAGCTCGCAGAGCCCAAGTTGC	1470	
Db	781	ACCAAGTCATATACAGCTTCAAAAGAGCTCTTGAGCTGTGAGCTCGCAGAGCCCAAGTTGC	840	
Qy	1471	ATACCACTGCCATTTGACAAAGAGGCTGTGGGCTGTAAAGCGGAGCTTATGCTGCTT	1530	
Db	841	ATACCACTGCCATTTGACAAAGAGGCTGTGGGCTGTAAAGCGGAGCTTATGCTGCTT	900	
Qy	1531	GCCAGGAAGCAAAAGAGATATAAAGAGTCAATTCAGAAAAATGTCTCTCAACATCCACTTC	1590	
Db	901	GCCAGGAAGCAAAAGAGATATAAAGAGTCAATTCAGAAAAATGTCTCTCAACATCCACTTC	960	
Qy	1591	ATGTAGAAGTATTTACCTCAGAGATTTATGGCTCATCAGAAAAATTTGCTTGGTCTTGCTT	1650	
Db	961	ATGTAGAAGTATTTACCTCAGAGATTTATGGCTCATCAGAAAAATTTGCTTGGTCTTGCTT	1020	
Qy	1651	CCTGGATGAACAAAATTTATGAGCTATTCAAGTGAATTTAGGCAGATCTTTTGGCCAAAGCAT	1710	
Db	1021	CCTGGATGAACAAAATTTATGAGCTATTCAAGTGAATTTAGGCAGATCTTTTGGCCAAAGCAT	1080	
Qy	1711	GCCTTAGAGAGAACCTGACTCGGAGAAATCCCTGTCTCATAGCAGGTTAATGCTTTGGG	1770	
Db	1081	GCCTTAGAGAGAACCTGACTCGGAGAAATCCCTGTCTCATAGCAGGTTAATGCTTTGGG	1140	
Qy	1771	ATGCAAGCTTTTAAAGGTGCGGAGTAAAGTTCCTTCATGAATTTGATCTTCAGCAGTTTTT	1830	
Db	1141	ATGCAAGCTTTTAAAGGTGCGGAGTAAAGTTCCTTCATGAATTTGATCTTCAGCAGTTTTT	1200	

1831 QY TTATGAGATGAATACAAAAAATCTTTGCTATGGAATTTGTGAAGTATTATAACAC 1890
1201 Db TTATGAGATGAATACAAAAAATCTTTGCTATGGAATTTGTGAAGTATTATAACAC 1260
1891 QY TGCAGAAAGATATATACGTGATGATCATGACAGAGTATCTATAAATGCACTTTTCA 1950
1261 Db TGCAGAAAGATATATACGTGATGATCATGACAGAGTATCTATAAATGCACTTTCA 1320
1951 QY TTACAGATTTTACTGTTCTTACTCTGGCTGCGACATCTTATGAGAGCAGATGTTATCT 2010
1321 Db TTACAGATTTTACTGTTCTTACTCTGGCTGCGACATCTTATGAGAGCAGATGTTATCT 1380
2011 QY CTGTCAATTAAGTAACTCTGCTAGAGTATTTTACCTGAGTACTTGGACAGAAATAAAT 2070
1381 Db CTGTCAATTAAGTAACTCTGCTAGAGTATTTTACCTGAGTACTTGGACAGAAATAAAT 1440
2071 QY TCAACTTCCAGGGTTATATCCAGGACAAATTTGGAGAGTATATGCAATATGTCACC 2130
1441 Db TCAACTTCCAGGGTTATATCCAGGACAAATTTGGAGAGTATATGCAATATGTCACC 1500
2131 QY TAAAGTATATCTGATCAGCAAAACCAATATGAGCAGAAAGATTAAGAATGCAATTC 2190
1501 Db TAAAGTATATCTGATCAGCAAAACCAATATGAGCAGAAAGATTAAGAATGCAATTC 1560
2191 QY TTGAGAGTTTTCGATCTTTTGAAGATTTTACCTGATGAGGGAATGGAAGAAATCC 2250
1561 Db TTGAGAGTTTTCGATCTTTTGAAGATTTTACCTGATGAGGGAATGGAAGAAATCC 1620
2251 QY GAAGCAGGTTGGGCAACACATTTGAAGTGGATCTTGATTTGGAGGCTGCAATTTGCTATAC 2310
1621 Db GAAGCAGGTTGGGCAACACATTTGAAGTGGATCTTGATTTGGAGGCTGCAATTTGCTATAC 1680
2311 QY AGATGCAATTTGAAGATATTTTACTCATGTTTCAAGAGTGGTGTCTTGTGATGAGAAC 2370
1681 Db AGATGCAATTTGAAGATATTTTACTCATGTTTCAAGAGTGGTGTCTTGTGATGAGAAC 1740
2371 QY TCTTACTTGTGCTTATTAAGATGTCACAAAGCTGTGATGAGTGCAGTACAGTTTCA 2430
1741 Db TCTTACTTGTGCTTATTAAGATGTCACAAAGCTGTGATGAGTGCAGTACAGTTTCA 1800
2431 QY TATCTAGTAGCAAGCAGTAGTACAAATCGTGTGGACATAGTTTGGAAACAAAGTCTTACA 2490
1801 Db TATCTAGTAGCAAGCAGTAGTACAAATCGTGTGGACATAGTTTGGAAACAAAGTCTTACA 1860
2491 QY GAGTATCTGAGGATCTTGTAAAGCATACTGTCACACTCTCTAGGACCTTGTGCTTTC 2550
1861 Db GAGTATCTGAGGATCTTGTAAAGCATACTGTCACACTCTCTAGGACCTTGTGCTTTC 1920
2551 QY ATGTAGGTTTAAAGCAGGCTGGTGTCTGTTTCAAGACTGCAATGATGTTGTCTTTTGG 2610
1921 Db ATGTAGGTTTAAAGCAGGCTGGTGTCTGTTTCAAGACTGCAATGATGTTGTCTTTTGG 1980
2611 QY ACTTTCAAGTAGAGTACTAGTGGATATCTTTTACGTTGCTGTTGTTGGTCCGAG 2670
1981 Db ACTTTCAAGTAGAGTACTAGTGGATATCTTTTACGTTGCTGTTGTTGGTCCGAG 2040
2671 QY TTGTGCTGAGATGTGGGCAAGAAATGGAATGCTCTTATAGCAGAGTGTTTTATACC 2730
2041 Db TTGTGCTGAGATGTGGGCAAGAAATGGAATGCTCTTATAGCAGAGTGTTTTATACC 2100
2731 QY AAGATGTTAAGTGACAGAGAAATGATGATAAAGATATCATATGTTTCAAGATGTTG 2790
2101 Db AAGATGTTAAGTGACAGAGAAATGATGATAAAGATATCATATGTTTCAAGATGTTG 2160
2791 QY CATCTTTAATGATCCCAATAGTTCTTGTGTTACTGTTACTTACAGAGTATGAACTTGG 2850
2161 Db CATCTTTAATGATCCCAATAGTTCTTGTGTTACTGTTACTTACAGAGTATGAACTTGG 2220
2851 QY AGGCTTTTAAAGACCAATATCTACAAAGACAGAGTTTGTATTAACAAATATATACAC 2910
2221 Db AGGCTTTTAAAGACCAATATCTACAAAGACAGAGTTTGTATTAACAAATATATACAC 2280
2911 QY TAATAGAGAAATGCTTCAGGTCCTCATCTATATTTGTGGGTGAGCGTTATGTCCTGGAG 2970

2281 Db TAATAGAGAAATGCTTCAGGTCCTCATCTATATTTGTGGGTGAGCGTTATGTACCTGGAG 2340
2971 QY TGGGAAATGTGACCAAGAGAGGTCACAATGAGAGAAATCAATTCACCTTGTTCATTTG 3030
2341 Db TGGGAAATGTGACCAAGAGAGGTCACAATGAGAGAAATCAATTCACCTTGTTCATTTG 2400
3031 QY AACCATGCGACACACAGTGCATTTGCCAAAAATTTTACCTGAGAAATGAATAATGAAC 3090
2401 Db AACCATGCGACACACAGTGCATTTGCCAAAAATTTTACCTGAGAAATGAATAATGAAC 2460
3091 QY GCTTAGAGAAATGCTATAAACCAAGTGGCCACATTTTAAGAAACCAAGTGTATCAGGCCATG 3150
2461 Db GCTTAGAGAAATGCTATAAACCAAGTGGCCACATTTTAAGAAACCAAGTGTATCAGGCCATG 2520
3151 QY GAGTTTATGAATCTAAAGATGAATCACTGAAGAGCTTCAATATGTACTTTTATCATTTACT 3210
2521 Db GAGTTTATGAATCTAAAGATGAATCACTGAAGAGCTTCAATATGTACTTTTATCATTTACT 2580
3211 QY CCAAAACCCAGCATAGCAAGGCTGAACATATGTCAGAGAAAGAGAGAAACAAAGAAACA 3270
2581 Db CCAAAACCCAGCATAGCAAGGCTGAACATATGTCAGAGAAAGAGAGAAACAAAGAAACA 2640
3271 QY AAGATGAAGCATTTGCGGCCACACCACTCTGTAATTTCTGCCCTGCTTTTACGAAAGTGA 3330
2641 Db AAGATGAAGCATTTGCGGCCACACCACTCTGTAATTTCTGCCCTGCTTTTACGAAAGTGA 2700
3331 QY TTAACTTCTCAACTGTGATATCATGATGTATCATTTCTCAGGACCGTATTTAGCGGGCAA 3390
2701 Db TTAACTTCTCAACTGTGATATCATGATGTATCATTTCTCAGGACCGTATTTAGCGGGCAA 2760
3391 QY TAGACACAGATTTCTAACTTTGTGGACCGAAGGATGCTCCAAATGCTTTTTCATATTTGG 3450
2761 Db TAGACACAGATTTCTAACTTTGTGGACCGAAGGATGCTCCAAATGCTTTTTCATATTTGG 2820
3451 QY CATTTGGGTTTACTAGAGAGAGCAACAGCTTCAAAAGCTCTTGAAGAGAGTAAACAT 3510
2821 Db CATTTGGGTTTACTAGAGAGAGCAACAGCTTCAAAAGCTCTTGAAGAGAGTAAACAT 2880
3511 QY TTGACTTTTATCAATAAGGCTTCAAGATTTGGGAAGTTTCCAGCATGAATATACAAATGCTTT 3570
2881 Db TTGACTTTTATCAATAAGGCTTCAAGATTTGGGAAGTTTCCAGCATGAATATACAAATGCTTT 2940
3571 QY TGGAAAACTCAAAGGAATTTCCAGTTAGAGGCGCAGAGGACATGATTAACGTCGATAC 3630
2941 Db TGGAAAACTCAAAGGAATTTCCAGTTAGAGGCGCAGAGGACATGATTAACGTCGATAC 3000
3631 QY TTCAGATGTTTGAACAGTGAAGGATTTAAGAGAAATCTTTGTTAATTTGTAGCAACCA 3690
3001 Db TTCAGATGTTTGAACAGTGAAGGATTTAAGAGAAATCTTTGTTAATTTGTAGCAACCA 3060
3691 QY CATCAGGATCGGAATCTTATTAAGATGATGAGATTTACTCATGATAAGAAAGAGCAAGAC 3750
3061 Db CATCAGGATCGGAATCTTATTAAGATGATGAGATTTACTCATGATAAGAAAGAGCAAGAC 3120
3751 QY GAAAAAGAAAAGCTGAGCTGCTAGGCTACATCGCCAGAGATCATGGCTCAGATGTCG 3810
3121 Db GAAAAAGAAAAGCTGAGCTGCTAGGCTACATCGCCAGAGATCATGGCTCAGATGTCG 3180
3811 QY CCTTACAGAAAACTTCAATTTGAAAATCTATAAATCTCATGTTATGACAAATCATCAGAAATGC 3870
3181 Db CCTTACAGAAAACTTCAATTTGAAAATCTATAAATCTCATGTTATGACAAATCATCAGAAATGC 3240
3871 QY CTGGGAAGAGATTTCCATTTATGGGAGAGAGACCCCGAGAGTACTGCTACTCTA 3930
3241 Db CTGGGAAGAGATTTCCATTTATGGGAGAGAGACCCCGAGAGTACTGCTACTCTA 3300
3931 QY GAATTTGCTTTGGGCTCTAAAACGGGCTCATCTGTTTACTGAAAAGAGGCTGCTGACGTCGA 3990
3301 Db GAATTTGCTTTGGGCTCTAAAACGGGCTCATCTGTTTACTGAAAAGAGGCTGCTGACGTCGA 3360
3991 QY TCCTTTGCGAAGAGAACAGAGGTTGAAAATATAGAAAATATAGCCATGTTATTCGGCT 4050

Db 3361 TCCTTTGCCAAGAGAACAGGAGGTGAAAAATAGAAAAATAATGCCATGGTATTATCGGCT 3420
Qy 4051 GTGTCCAGAAATCTACTGCTTAACCCAGCAGCAGGGGAAAAACCCATAGAACTCTCAGGAG 4110
Db 3421 GTGTCCAGAAATCTACTGCTTAACCCAGCAGCAGGGGAAAAACCCATAGAACTCTCAGGAG 3480
Qy 4111 AAGCCCTAGAACCCACTTTTCATGGATCAGAGCTTGGCATATGGAACTTATACAGGAAGCT 4170
Db 3481 AAGCCCTAGAACCCACTTTTCATGGATCAGAGCTTGGCATATGGAACTTATACAGGAAGCT 3540
Qy 4171 GTGTGATGTAAATGACAGCAGTGTCTGGCAGAAATATTTTGAAGCTGTACAGCTGAGCT 4230
Db 3541 GTGTGATGTAAATGACAGCAGTGTCTGGCAGAAATATTTTGAAGCTGTACAGCTGAGCT 3600
Qy 4231 CTCAGCAGCGCATTCATGTGACCTTTTGTGACTTTGGAAAGTGGAAATATCTTTGCCCTC 4290
Db 3601 CTCAGCAGCGCATTCATGTGACCTTTTGTGACTTTGGAAAGTGGAAATATCTTTGCCCTC 3660
Qy 4291 TTTTGCATAATCTGTGCAATATCTGTGATCCCATTTATTTCTTTGCAACTCAAAAGATAA 4350
Db 3661 TTTTGCATAATCTGTGCAATATCTGTGATCCCATTTATTTCTTTGCAACTCAAAAGATAA 3720
Qy 4351 ACAGTGAGATGCGAGATGCTCTTGCTCAACTTTTGAACCTTGGCAGCGTGGATACAGACTG 4410
Db 3721 ACAGTGAGATGCGAGATGCTCTTGCTCAACTTTTGAACCTTGGCAGCGTGGATACAGACTG 3780
Qy 4411 TTCTGGCCAGAAATATCAGGTTATATATAGACATGCTTAAAGGAGAAACCCAAATTCCTA 4470
Db 3781 TTCTGGCCAGAAATATCAGGTTATATATAGACATGCTTAAAGGAGAAACCCAAATTCCTA 3840
Qy 4471 TTTTCTTTTAAACAGGAATGSGAGATCTACTTTGGAGTCCATTTCCATCTGAGTTTTG 4530
Db 3841 TTTTCTTTTAAACAGGAATGSGAGATCTACTTTGGAGTCCATTTCCATCTGAGTTTTG 3900
Qy 4531 GCGTTGAGTCTTCGATTAATAATCAAAATAGCATCAAGGAAATGGTTATTTCTTTTGCCA 4590
Db 3901 GCGTTGAGTCTTCGATTAATAATCAAAATAGCATCAAGGAAATGGTTATTTCTTTTGCCA 3960
Qy 4591 CAACAAATTTATAGAAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATG 4650
Db 3961 CAACAAATTTATAGAAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATG 4020
Qy 4651 TGCTGACCTTGAGACCTGCGCTTTCACTATCCAGGCAATGAAATCTATTGGGAGATG 4710
Db 4021 TGCTGACCTTGAGACCTGCGCTTTCACTATCCAGGCAATGAAATCTATTGGGAGATG 4080
Qy 4711 AAGGAAAACTCTGTTTGGAGCACTTCAAAATFAGGCAGCATAAATGGTCTGAAAGCATTA 4770
Db 4081 AAGGAAAACTCTGTTTGGAGCACTTCAAAATFAGGCAGCATAAATGGTCTGAAAGCATTA 4140
Qy 4771 TGCAGTTTGCAGTTGCAAGAGATTAATCTGTCTCAGGCTCTGATACAGAAACATCTGG 4830
Db 4141 TGCAGTTTGCAGTTGCAAGAGATTAATCTGTCTCAGGCTCTGATACAGAAACATCTGG 4200
Qy 4831 TTGCTCTTCTACAGTGTCTTCTTCACTAATAAATCAGAGATACACCATGCTTCTGT 4890
Db 4201 TTGCTCTTCTACAGTGTCTTCTTCACTAATAAATCAGAGATACACCATGCTTCTGT 4260
Qy 4891 CTATAGATCTGTTTCAATGTTTGGTGGTGTGTGTAGCATTCCTCATCTTGTATTGGG 4950
Db 4261 CTATAGATCTGTTTCAATGTTTGGTGGTGTGTGTAGCATTCCTCATCTTGTATTGGG 4320
Qy 4951 ATGACCTTGTGATCTGACGCTTCTTCAAGTTAGTTCCTTATTAACCACTTTATCTCT 5010
Db 4321 ATGACCTTGTGATCTGACGCTTCTTCAAGTTAGTTCCTTATTAACCACTTTATCTCT 4380
Qy 5011 TCCATTTGATCACCATGACACATGCTTCAGATACTACTTACAGTACACAGCCCTAC 5070
Db 4381 TCCATTTGATCACCATGACACATGCTTCAGATACTACTTACAGTACACAGCCCTAC 4440
Qy 5071 CCCTTGCTCAGGTTCAAGAGACAGTGAAGGCTCATTTCCGCTCTTTCTTTTGAG 5130
Db 4441 CCCTTGCTCAGGTTCAAGAGACAGTGAAGGCTCATTTCCGCTCTTTCTTTTGAG 4500

Qy 5131 AAATTTCTCAATATACAAGTGGCTCCATTTGGTGTGATATTTCTGGCTGGTATTTGTGGG 5190
Db 4501 AAATTTCTCAATATACAAGTGGCTCCATTTGGTGTGATATTTCTGGCTGGTATTTGTGGG 4560
Qy 5191 TCTCACTGAAGAAATGGCATCACCCCTTATCTTCGCTGTGCTGCAATTTGTTTTCACATAT 5250
Db 4561 TCTCACTGAAGAAATGGCATCACCCCTTATCTTCGCTGTGCTGCAATTTGTTTTCACATAT 4620
Qy 5251 TACTTGGGTTAACTCCGCTGAGGAACCTGCATACCAATTTCTGCAGAGGAGGTACAGTG 5310
Db 4621 TACTTGGGTTAACTCCGCTGAGGAACCTGCATACCAATTTCTGCAGAGGAGGTACAGTG 4680
Qy 5311 CACTCTGTAGCTATCTATCTTTTACTCAAAATTTGTTCTGCTCTTCCAGGAATATTGGG 5370
Db 4681 CACTCTGTAGCTATCTATCTTTTACTCAAAATTTGTTCTGCTCTTCCAGGAATATTGGG 4740
Qy 5371 ATACTGTAAGGCCCTTCTCCAGAGTGTGTGAGATCTTGCCTTACTAACTGTTTGA 5430
Db 4741 ATACTGTAAGGCCCTTCTCCAGAGTGTGTGAGATCTTGCCTTACTAACTGTTTGA 4800
Qy 5431 AGCAAAAAACACCGTGTGAGTACCCTAGAAAAAGAAATAGTTTGTATAGAGCTTCTG 5490
Db 4801 AGCAAAAAACACCGTGTGAGTACCCTAGAAAAAGAAATAGTTTGTATAGAGCTTCTG 4860
Qy 5491 ATGACTATAGCTGCTCTCTGAATCAAGCTTCTCAATTTCCAGTGCCTCAGGCTCAGATG 5550
Db 4861 ATGACTATAGCTGCTCTCTGAATCAAGCTTCTCAATTTCCAGTGCCTCAGGCTCAGATG 4920
Qy 5551 ATGAGCGAAAGCATCTCTCTGCTCTTCTGCGGCTTATCTATGTTTCTCAGAAC 5610
Db 4921 ATGAGCGAAAGCATCTCTCTGCTCTTCTGCGGCTTATCTATGTTTCTCAGAAC 4980
Qy 5611 TTTTCTGCGAGAAATTTGTGAACGCGGGAAGGTTTGGAGCTTGCAATTTTTCACGCACTTC 5670
Db 4981 TTTTCTGCGAGAAATTTGTGAACGCGGGAAGGTTTGGAGCTTGCAATTTTTCACGCACTTC 5040
Qy 5671 ACTGTGAGCGGAGTCTGCAATTTTCTTAAATTCAGAGATGCGAGTGTCTTGGTGG 5730
Db 5041 ACTGTGAGCGGAGTCTGCAATTTTCTTAAATTCAGAGATGCGAGTGTCTTGGTGG 5100
Qy 5731 AAGTAAAGCCAGAGGCTGTGCTTATCCAGCTCTTCTTCTTGGATGAATATGAGAAACAG 5790
Db 5101 AAGTAAAGCCAGAGGCTGTGCTTATCCAGCTCTTCTTCTTGGATGAATATGAGAAACAG 5160
Qy 5791 ACCCTGGCTGAGAGGGGCAACCCCTTCAATTTATCTGTGAGGGTATCGAGCTCC 5850
Db 5161 ACCCTGGCTGAGAGGGGCAACCCCTTCAATTTATCTGTGAGGGTATCGAGCTCC 5220
Qy 5851 ATTTGCTGTGGCAACACCTGCAATTTATAGAGAGATTTGCTAGGAGCCAGAGACTAATC 5910
Db 5221 ATTTGCTGTGGCAACACCTGCAATTTATAGAGAGATTTGCTAGGAGCCAGAGACTAATC 5280
Qy 5911 AGATGTTATTTGGATTCAACTGGCAGTTTACTGTGAGCTCCAACTCTGCTCAAGCAATC 5970
Db 5281 AGATGTTATTTGGATTCAACTGGCAGTTTACTGTGAGCTCCAACTCTGCTCAAGCAATC 5340
Qy 5971 ACAAATGACGACAGTAGTAAGGCTGATCAAAATTTATGAAAACTTCTGAGGGCTGGG 6030
Db 5341 ACAAATGACGACAGTAGTAAGGCTGATCAAAATTTATGAAAACTTCTGAGGGCTGGG 5400
Qy 6031 AAAGTATTTGGAGGCTCTTTTGTCTCCATGCTCAGGTTCACTTTACATCAATAAATATTTCT 6090
Db 5401 AAAGTATTTGGAGGCTCTTTTGTCTCCATGCTCAGGTTCACTTTACATCAATAAATATTTCT 5460
Qy 6091 TAATGG 6096
Db 5461 TAATGG 5466

```
; Publication No. US20040185037A1
; GENERAL INFORMATION:
; APPLICANT: Han, et al.
; TITLE OF INVENTION: HUMAN E3 ALPHA UBIQUITIN LIGASE FAMILY
; FILE REFERENCE: 01017/35966B
; CURRENT APPLICATION NUMBER: US/10/758,672A
; CURRENT FILING DATE: 2004-01-15
; PRIOR APPLICATION NUMBER: US 09/724,126
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 60/187,911
; PRIOR FILING DATE: 2000-03-08
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18
; LENGTH: 5205
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (662)..(662)
; OTHER INFORMATION: n = a, c, g, or t
; NAME/KEY: misc feature
; LOCATION: (668)..(668)
; OTHER INFORMATION: n = a, c, g, or t
; US-10-758-672A-18

Query Match      81.4%; Score 5136; DB 19; Length 5205;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 5189; Conservative 10; Mismatches 6; Indels 45; Gaps 1;

QY 696 ATGGCGGACGAGGAGGCTGGAGTACTGAGAGGATGGAATCAGCGCGAGGTACCCAG 755
DB 1 ATGGCGGACGAGGAGGCTGGAGTACTGAGAGGATGGAATCAGCGCGAGGTACCCAG 60
QY 756 ACCCTCAGCGCTCGCACTCTGGGGATCAGCAAGTTGATTTTATCTGCTTTCTTG 815
DB 61 ACCCTCAGCGCTCGCACTCTGGGGATCAGCAAGTTGATTTTATCTGCTTTCTTG 120
QY 816 CATCATTTGGCAATTTGGTGCAGAAATTTACTTTGCTGGAATGGACCCAGACTTGGAA 875
DB 121 CATCATTTGGCAATTTGGTGCAGAAATTTACTTTGCTGGAATGGACCCAGACTTGGAA 180
QY 876 AAGCAGGAGGAAAGTGTAACAATGTCAATATTCATCCACTGGAATGTACTTATTGGA 935
DB 181 AAGCAGGAGGAAAGTGTAACAATGTCAATATTCATCCACTGGAATGTACTTATTGGA 240
QY 936 GAAGATCCAGATATTTGCTTAGAGAAATTTAGACACAGTGGAGCAATTTACGTTTGTGG 995
DB 241 GAAGATCCAGATATTTGCTTAGAGAAATTTAGAGCAAGTGGAGCAATTTACGTTTGTGG 300
QY 996 AGGGTTTTCAAAGTGGAGAGACCACTTATTTCTGAGGATTTGTCAATTTGATCCAACA 1055
DB 301 AGGGTTTTCAAAGTGGAGAGACCACTTATTTCTGAGGATTTGTCAATTTGATCCAACA 360
QY 1056 TGTGTACTCTGTATGAGCTGCTTCAGACAGTGTTCATAAAATTCATCGTTACAGATG 1115
DB 361 TGTGTACTCTGTATGAGCTGCTTCAGACAGTGTTCATAAAATTCATCGTTACAGATG 420
QY 1116 CATACTTCTACTGGAGGAGGTTCTGTGACTGTGTGAGACACAGAGGCAATGGAATCTGGC 1175
DB 421 CATACTTCTACTGGAGGAGGTTCTGTGACTGTGTGAGACACAGAGGCAATGGAATCTGGC 480
QY 1176 CCTTTTCTGTAAATCATGAACCTGGAAGAGCAGGTACTATATAAGAGAAATTCAGCTGT 1235
DB 481 CCTTTTCTGTAAATCATGAACCTGGAAGAGCAGGTACTATATAAGAGAAATTCAGCTGT 540
QY 1236 CCGTTGAATGAAGAGGTAATTTGTCAGCCAGGAAATATTTTCTTCTAGTGATAAATAT 1295
DB 541 CCGTTGAATGAAGAGGTAATTTGTCAGCCAGGAAATATTTTCTTCTAGTGATAAATAT 600
QY 1296 GTCGTAGAAATGACTATATGGAGAGGAAAGAACTGCGCTCTCTGAACTCCAGATAAGG 1355
```


QY 2436 AGTAGAAGACAGTAGTACAATCGTGTGGACATAGTCTTTGGAAAAAACAAGTCTCTACAGAGTA 2495
Db
QY 1696 AGTAGAAGACAGTAGTACAATCGTGTGGACATAGTCTTTGGAAAAAACAAGTCTCTACAGAGTA 1755
Db
QY 2496 TCTGAGGATCTTTGTAAGCATACATCTGCCACTCTCTAGGACCCCTTGCTGGTCTTTCATGTA 2555
Db
QY 1756 TCTGAGGATCTTTGTAAGCATACATCTGCCACTCTCTAGGACCCCTTGCTGGTCTTTCATGTA 1815
QY 2556 CGTTTAAGCAGGCTGGGTGCTGTTTCAAGACCTGCATGAAATTTGTGTCTTTTGAGGACTTTT 2615
Db
QY 1816 CGTTTAAGCAGGCTGGGTGCTGTTTCAAGACCTGCATGAAATTTGTGTCTTTTGAGGACTTTT 1875
QY 2616 CAAGTAGAGGTACTAGTGGAAATATCCTTTTACGTTGCTCTGGTGTGGTGGCCAGGTTCTT 2675
Db
QY 1876 CAAGTAGAGGTACTAGTGGAAATATCCTTTTACGTTGCTCTGGTGTGGTGGCCAGGTTGTT 1935
QY 2676 GCTGAGATGTGGCGAAGAAATGGAGCTGTCTCTTATTAGCCAGGTGTTTATTATCAAGAT 2735
Db
QY 1936 GCTGAGATGTGGCGAAGAAATGGAGCTGTCTCTTATTAGCCAGGTGTTTATTATCAAGAT 1995
QY 2736 GTTAAGTGCAGAGAAATGTATGATTAAGATATCATGCTTTCAGATTTGGTGCATCT 2795
Db
QY 1996 GTTAAGTGCAGAGAAATGTATGATTAAGATATCATGCTTTCAGATTTGGTGCATCT 2055
QY 2796 TTAAATGGATCCCAATAAGTCTTGTGTTACTGTTACTCTCAGAGGTATGAACCTTGGCCGAGGCT 2855
Db
QY 2056 TTAAATGGATCCCAATAAGTCTTGTGTTACTGTTACTCTCAGAGGTATGAACCTTGGCCGAGGCT 2115
QY 2856 TTAAACAAGACCATATCTACAAAAGACAGGATTTGATTAACAATAATAATACACTAATA 2915
Db
QY 2116 TTAAACAAGACCATATCTACAAAAGACAGGATTTGATTAACAATAATAATACACTAATA 2175
QY 2916 GAAGAAATGCTTCAGGTCTCTATCTATATTTGGGTGAGCGTATGTAACCTGGAGTGGGA 2975
Db
QY 2176 GAAGAAATGCTTCAGGTCTCTATCTATATTTGGGTGAGCGTATGTAACCTGGAGTGGGA 2235
QY 2976 AATGTGACCAAGAGAGGTCAATGAGAGAAATCAATTCATCTGCTTTGCAATGGAACCC 3035
Db
QY 2236 AATGTGACCAAGAGAGGTCAATGAGAGAAATCAATTCATCTGCTTTGCAATGGAACCC 2295
QY 3036 ATGCCACACAGTGCATTTGCCAAAATTTTACCTGAGAAATGAAAATAATGAAACTGGCTTA 3095
Db
QY 2296 ATGCCACACAGTGCATTTGCCAAAATTTTACCTGAGAAATGAAAATAATGAAACTGGCTTA 2355
QY 3096 GAGAAATGTCAATAACAAGTGGCCACATTTAAGAAACCAAGTGTATCAGGCCATGGAGTT 3155
Db
QY 2356 GAGAAATGTCAATAACAAGTGGCCACATTTAAGAAACCAAGTGTATCAGGCCATGGAGTT 2415
QY 3156 TATGAACTTAAGAGATGAAATCACTGAAAGACTTCAATATGTACTTTTATCATTTACTCCAAA 3215
Db
QY 2416 TATGAACTTAAGAGATGAAATCACTGAAAGACTTCAATATGTACTTTTATCATTTACTCCAAA 2475
QY 3216 ACCCAGCATAGCAAGGCTGAAACATATGAGAGAAAGAGAGAAACAAAGAT 3275
Db
QY 2476 ACCCAGCATAGCAAGGCTGAAACATATGAGAGAAAGAGAGAAACAAAGAT 2535
QY 3276 GAAGCAATTTGCCCCACCAACCTCTCTGAAATTTCTGCCCTGCTTTCAGCAAAAGTGAATTAAC 3335
Db
QY 2536 GAAGCAATTTGCCCCACCAACCTCTCTGAAATTTCTGCCCTGCTTTCAGCAAAAGTGAATTAAC 2595
QY 3336 CTTCTCAACTGTGATATCATGATGTACATTTCTCAGGACCGTATTTTGGGGGCAATAGAC 3395
Db
QY 2596 CTTCTCAACTGTGATATCATGATGTACATTTCTCAGGACCGTATTTTGGGGGCAATAGAC 2655
QY 3396 ACAGATTTCTAACTGTGACCGAAGGATGCTCCAAATGGCTTTTCAATATCTGGCATTTG 3455
Db
QY 2656 ACAGATTTCTAACTGTGACCGAAGGATGCTCCAAATGGCTTTTCAATATCTGGCATTTG 2715
QY 3456 GGTTTTACTAGAAAGAGCAACAGCTTCAAAAAGCTTCTGAAAGAAAGATTAACATTTGAC 3515
Db
QY 2716 GGTTTTACTAGAAAGAGCAACAGCTTCAAAAAGCTTCTGAAAGAAAGATTAACATTTGAC 2775

QY 3516 TTTTATCATAAAGCTTCAAGATTTGGGAAGTTTACGCCATGAATATACAAATGCTTTTGGAA 3575
Db
QY 2776 TTTTATCATAAAGCTTCAAGATTTGGGAAGTTTACGCCATGAATATACAAATGCTTTTGGAA 2835
Db
QY 3576 AAACCTCAAGGAATTTCCCAAGTTTGAAGGCCAGAGGACATGATAAAGTGTGATTTCTCAG 3635
Db
QY 2836 AAACCTCAAGGAATTTCCCAAGTTTGAAGGCCAGAGGACATGATAAAGTGTGATTTCTCAG 2895
QY 3636 ATGTTTTCACACAGTGAAGCGATTAAGAGAAATAATCTTTGTTTAAATTTGTAGCAACCAATCA 3695
Db
QY 2896 ATGTTTTCACACAGTGAAGCGATTAAGAGAAATAATCTTTGTTTAAATTTGTAGCAACCAATCA 2955
QY 3696 GGATCGGAATCTTAAAGAAATGATGAGATTAATCATGATAAAGAAAAGCAGAAACGAAAA 3755
Db
QY 2956 GGATCGGAATCTTAAAGAAATGATGAGATTAATCATGATAAAGAAAAGCAGAAACGAAAA 3015
QY 3756 AGAAAAGCTGAAGCTGTAGGCTACATCGCCAGAGATCATGGCTCAGATGTCTGCCCTTA 3815
Db
QY 3016 AGAAAAGCTGAAGCTGTAGGCTACATCGCCAGAGATCATGGCTCAGATGTCTGCCCTTA 3075
QY 3816 CAGAAAACCTTCAATTTGAAAACCTCATAAACTCATGTATGACAATACATCAGAAATGCTCGGG 3875
Db
QY 3076 CAGAAAACCTTCAATTTGAAAACCTCATAAACTCATGTATGACAATACATCAGAAATGCTCGGG 3135
QY 3876 AAAGAAGATTCATTTATGGAGGAAGAGAGCACCCAGCAGTCACTAGTAATTT 3935
Db
QY 3136 AAAGAAGATTCATTTATGGAGGAAGAGAGCACCCAGCAGTCACTAGTAATTT 3195
QY 3936 GCTTTGGGTCTTAAACGGGGTCCATCTGTTTACTGAAAAGAGGAGTGTGACGTGATCTT 3995
Db
QY 3196 GCTTTGGGTCTTAAACGGGGTCCATCTGTTTACTGAAAAGAGGAGTGTGACGTGATCTT 3255
QY 3996 TGCCAAAGAGAACAGGAGGTGAAAATAAGAAAATAATGCAATGTTATTTATCGGCTGTGTC 4055
Db
QY 3256 TGCCAAAGAGAACAGGAGGTGAAAATAAGAAAATAATGCAATGTTATTTATCGGCTGTGTC 3315
QY 4056 CAGAAAATCTACTGCTCTTAAACCCAGCACAGGGGAAAAACCCATAGAACTCTCAGGAGAAAGCC 4115
Db
QY 3316 CAGAAAATCTACTGCTCTTAAACCCAGCACAGGGGAAAAACCCATAGAACTCTCAGGAGAAAGCC 3375
QY 4116 CTAGACCCCATCTTTTCTAGATCCAGATCTGGCATATGGAACTTATACAGGAAGCTGTGTT 4175
Db
QY 3376 CTAGACCCCATCTTTTCTAGATCCAGATCTGGCATATGGAACTTATACAGGAAGCTGTGTT 3435
QY 4176 CATGTATGACGACGATGTGCTGACAGAGTATTTTGAAGCTGTACAGCTCAGCTCTCAG 4235
Db
QY 3436 CATGTATGACGACGATGTGCTGACAGAGTATTTTGAAGCTGTACAGCTCAGCTCTCAG 3495
QY 4236 CAGCGCATTTCAATGTTGACCTTTTGTGACCTTGGAAAAGTGGAGAAATATCTTTGCCCTCTTTGC 4295
Db
QY 3496 CAGCGCATTTCAATGTTGACCTTTTGTGACCTTGGAAAAGTGGAGAAATATCTTTGCCCTCTTTGC 3555
QY 4296 AAATCTCTGTGCAATATCTGTGATCCCATTTATTTCTTTGCAACCTCAAAAAGATAAAGAT 4355
Db
QY 3556 AAATCTCTGTGCAATATCTGTGATCCCATTTATTTCTTTGCAACCTCAAAAAGATAAAGAT 3615
QY 4356 GAGATGACAGATGCTCTTGTCTCAACTTTTGACCTTGCCACGCTGGATACAGACTGTTCTG 4415
Db
QY 3616 GAGATGACAGATGCTCTTGTCTCAACTTTTGACCTTGCCACGCTGGATACAGACTGTTCTG 3675
QY 4416 GCCAGAAATATCAGGTTTATAATATAAGACATGCTTAAAGGAGAAAAACCCAAATTTCTATTTTC 4475
Db
QY 3676 GCCAGAAATATCAGGTTTATAATATAAGACATGCTTAAAGGAGAAAAACCCAAATTTCTATTTTC 3735
QY 4476 TTTAATCAAGGAATGGGAGATTTCTATTTGGAGTTCCATTCATCTGAGTTTGGCGTT 4535
Db
QY 3736 TTTAATCAAGGAATGGGAGATTTCTATTTGGAGTTCCATTCATCTGAGTTTGGCGTT 3795
QY 4536 GAGTCTTCGATTTAAATATTTCAAAATAGCATCAAGGAATGTTTATTTCTTTGGCCACAACA 4595
Db
QY 3796 GAGTCTTCGATTTAAATATTTCAAAATAGCATCAAGGAATGTTTATTTCTTTGGCCACAACA 3855
QY 4596 ATTTATAGAAATTTGGAATTTGAAAGTCCACCTGATGAAAGGGATCTCTGAGTCCCATGCTG 4655

Db 3856 ATTTATAGAAATGGATGGAAGTGGACCTGATGAAGGATCTCGAGTCCCAATGCTG 3915
Qy 4656 ACCTGGAGCACTGGGCTTTCACTATCCAGGCAATTTGAAATCTATTGGGAGATCAAGGA 4715
Db 3916 ACCTGGAGCACTGGGCTTTCACTATCCAGGCAATTTGAAATCTATTGGGAGATGAAGA 3975
Qy 4716 AAACCTCTGTTTGGAGCACTTCAAAATAGGAGCAATAATGGTCTGAAAGCAATTAATGCA 4775
Db 3976 AAACCTCTGTTTGGAGCACTTCAAAATAGGAGCAATAATGGTCTGAAAGCAATTAATGCA 4035
Qy 4776 TTTCAGATGGACAGAGGATTAACCTGCTCAGGTCTGATACAGAAATCTGCTTCTGCTATA 4835
Db 4036 TTTCAGATGGACAGAGGATTAACCTGCTCAGGTCTGATACAGAAATCTGCTTCTGCTATA 4095
Qy 4836 TTTCATCAGATGTTTCTTCTTAACATAAAATCAGAGATACACCATGCTTCTGCTATA 4895
Db 4096 TTTCATCAGATGTTTCTTCTTAACATAAAATCAGAGATACACCATGCTTCTGCTATA 4155
Qy 4896 GATCTGTTTCATGTTTGGTGGTGTGTGTTAGCATTCCTTATCCCATCTTGTATGGAGTAC 4955
Db 4156 GATCTGTTTCATGTTTGGTGGTGTGTGTTAGCATTCCTTATCCCATCTTGTATGGAGTAC 4215
Qy 4956 CCGTGTGATCTGACGCTTCTTCACTAGTCTTCTTCTTATTAACCATCTTATCTTCCAT 5015
Db 4216 CCGTGTGATCTGACGCTTCTTCACTAGTCTTCTTCTTATTAACCATCTTATCTTCCAT 4275
Qy 5016 TTGATCACCATGGCACACATGCTTCAGATACCTTCTTACAGTACACAGGCTTACCCCTT 5075
Db 4276 TTGATCACCATGGCACACATGCTTCAGATACCTTCTTACAGTACACAGGCTTACCCCTT 4335
Qy 5076 GCTCAGGTTCAAGAGACAGTGAAGGCTCAATCCGATCTTCTTCTTCTTTCAGAAAT 5135
Db 4336 GCTCAGGTTCAAGAGACAGTGAAGGCTCAATCCGATCTTCTTCTTCTTTCAGAAAT 4395
Qy 5136 TCTCAATATACAGTGGCTCCATTTGGTGTGATATCTTGGCTGTGATTTTGGGTCTCA 5195
Db 4396 TCTCAATATACAGTGGCTCCATTTGGTGTGATATCTTGGCTGTGATTTTGGGTCTCA 4455
Qy 5196 CTGAAGATGGCATCACCCCTTATCTTCTGCTGTGCTGATTTCTTCTTCTTCTTCTT 5255
Db 4456 CTGAAGATGGCATCACCCCTTATCTTCTGCTGTGCTGATTTCTTCTTCTTCTTCTT 4515
Qy 5256 GGGGTAACTCCGCTGAGGAACTGCATACCAATCTGCAGAGGAGAGTACAGTCACTC 5315
Db 4516 GGGGTAACTCCGCTGAGGAACTGCATACCAATCTGCAGAGGAGAGTACAGTCACTC 4575
Qy 5316 TGTAGCTATCTATCTTTACCTACAAATTTGTTCTGCTCTTCCAGGAATATTTGGATACT 5375
Db 4576 TGTAGCTATCTATCTTTACCTACAAATTTGTTCTGCTCTTCCAGGAATATTTGGATACT 4635
Qy 5376 GTAAGGCTTGTCTCAGAGGTGGTGTGCAGATCTGCTTACTTAACTGTTTGAAGCAA 5435
Db 4636 GTAAGGCTTGTCTCAGAGGTGGTGTGCAGATCTGCTTACTTAACTGTTTGAAGCAA 4695
Qy 5436 AAAAACACCGTGGTCAAGGTACCTTAGAAAAAGAAATAGTTTGATAGAGCTTCTGATGAC 5495
Db 4696 AAAAACACCGTGGTCAAGGTACCTTAGAAAAAGAAATAGTTTGATAGAGCTTCTGATGAC 4755
Qy 5496 TATAGCTGCTCTGAAATCAAGCTTCTCATTTTCAAGTGGTCCAGCTTCTGAGATGATGAG 5555
Db 4756 TATAGCTGCTCTGAAATCAAGCTTCTCATTTTCAAGTGGTCCAGCTTCTGAGATGATGAG 4815
Qy 5556 CGAAGGATCTGCTCTGCTCTTCTGCTGCTTCTGCTGCTTCTGCTGCTTCTGCTGCTG 5615
Db 4816 CGAAGGATCTGCTCTGCTCTTCTGCTGCTTCTGCTGCTTCTGCTGCTTCTGCTGCT 4875
Qy 5616 TGCCAGGAAATTTGTAAGGGGAGAGGTTGGAGCTTGGATTTTTCAGGCACTTCACTGT 5675
Db 4876 TGCCAGGAAATTTGTAAGGGGAGAGGTTGGAGCTTGGATTTTTCAGGCACTTCACTGT 4935
Qy 5676 GGAGCCGAGTGTGATTTTCTTAAATCAGAGAAATCCGAGTGGTCTTGGTTGAAGGT 5735

Db 4936 GGAGCCGAGTGTGATTTTCTTAAATCAGAGAAATCCGAGTGGTCTTGGTTGAAGGT 4995
Qy 5736 AAAGCCAGAGGCTGTGCTTATCCAGCTCTTCTTATCTTGGATGAATATGGAGAAACAGACCT 5795
Db 4996 AAAGCCAGAGGCTGTGCTTATCCAGCTCTTCTTATCTTGGATGAATATGGAGAAACAGACCT 5055
Qy 5796 GGCCTGAAGAGGGGCAACCCCTTCTTATCTTCTGAGCGGTATCGGAAGCTCCTATTG 5855
Db 5056 GGCCTGAAGAGGGGCAACCCCTTCTTATCTTCTGAGCGGTATCGGAAGCTCCTATTG 5115
Qy 5856 GTCTGGCAACCACTGCTTATTAAGAGGATTTCTAGAGCCAGAGACTTAATCAGATG 5915
Db 5116 GTCTGGCAACCACTGCTTATTAAGAGGATTTCTAGAGCCAGAGACTTAATCAGATG 5175
Qy 5916 TTATTGGATTCAACTGGCAGTTACTGTGA 5945
Db 5176 TTATTGGATTCAACTGGCAGTTACTGTGA 5205

RESULT 7

US-10-758-636A-18
; Sequence 18, Application US/10758636A
; Publication No. US20050089876A1
; GENERAL INFORMATION:
; APPLICANT: Han, et al.
; TITLE OF INVENTION: HUMAN E3 ALPHA UBIQUITIN LIGASE FAMILY
; FILE REFERENCE: 01017/35966C
; CURRENT APPLICATION NUMBER: US/10758,636A
; PRIORITY FILING DATE: 2004-01-15
; PRIOR APPLICATION NUMBER: US 09/724,126
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 60/187,911
; PRIOR FILING DATE: 2000-03-08
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18
; LENGTH: 5205
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (662)..(662)
; OTHER INFORMATION: n = a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (668)..(668)
; OTHER INFORMATION: n = a, c, g, or t
US-10-758-636A-18

Query Match 81.4%; Score 5136; DB 21; Length 5205;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 5189; Conservative 10; Mismatches 6; Indels 45; Gaps 1;

Qy 696 ATGGCGGACGAGGAGGCTGAGGTAATCTGAGAGGATGGAATCAGCGCGAGGTACCCAG 755
Db 1 ATGGCGGACGAGGAGGCTGAGGTAATCTGAGAGGATGGAATCAGCGCGAGGTACCCAG 60
Qy 756 ACCCTCAGCTCTGGCATCTTGGTGGGATCAGCAAGTTGATTTTATCTGCTTCTTCTG 815
Db 61 ACCCTCAGCTCTGGCATCTTGGTGGGATCAGCAAGTTGATTTTATCTGCTTCTTCTG 120
Qy 816 CATCATTTGGCAAAATTTGGTCCAGAAATTTTCTGCTGAAATGGAAATCCAGACTTGGAA 875
Db 121 CATCATTTGGCAAAATTTGGTCCAGAAATTTTCTGCTGAAATGGAAATCCAGACTTGGAA 180
Qy 876 AAGCAGGAGGAAAGTGTACAAATGTCAATATTCATCCATCGGAATGTGATCTTATTGGA 935
Db 181 AAGCAGGAGGAAAGTGTACAAATGTCAATATTCATCCATCGGAATGTGATCTTATTGGA 240
Qy 936 GAAGATCCAGATATTTGCTTAGAGAAATTTGAAGCAGCAGTGGAGCATTTTCACTTTGTCGG 995
Db 241 GAAGATCCAGATATTTGCTTAGAGAAATTTGAAGCAGCAGTGGAGCATTTTCACTTTGTCGG 300

Db 2416 TATGAACATAAAGATGAATCACTGAAAGACCTTCAATATATGTACTTTTATCATTTACTCCAAA 2475
Qy 3216 ACCGAGCATPAGCAGGCTGAACATATGCGAAGAAAGAGAGAAAACAAGAAAACAAGAT 3275
Db 2476 ACCGAGCATPAGCAGGCTGAACATATGCGAAGAAAGAGAGAAAACAAGAAAACAAGAT 2535
Qy 3276 GAAGCATTTGCCGCCACCAACCACTCTGAAATTTCTGCAATTTCTGCGCTTTTTCAGCAGAGTGAATTAAC 3335
Db 2536 GAAGCATTTGCCGCCACCAACCACTCTGAAATTTCTGCAATTTCTGCGCTTTTTCAGCAGAGTGAATTAAC 2595
Qy 3336 CTTCCTCAACTGTGATATCATGATGTACATTTCTCAGACCGTATTTGAGCGGGCAATAGAC 3395
Db 2596 CTTCCTCAACTGTGATATCATGATGTACATTTCTCAGACCGTATTTGAGCGGGCAATAGAC 2655
Qy 3396 ACAGATTTCAACTGTGACCGAAGGGATGCTCCAAATGGCTTTTTCATATTTCTGCAATG 3455
Db 2656 ACAGATTTCAACTGTGACCGAAGGGATGCTCCAAATGGCTTTTTCATATTTCTGCAATG 2715
Qy 3456 GGTTTACTAGAGAGAGAGCAACGCTTCAAAAGCTTCCGAGAGAGAGTAAATTTGAC 3515
Db 2716 GGTTTACTAGAGAGAGAGCAACGCTTCAAAAGCTTCCGAGAGAGAGTAAATTTGAC 2775
Qy 3516 TTTTATCATTAAGGCTTCAAGATTGGCAAGTTCCAGCATGATATACAAATGCTTTTGGAA 3575
Db 2776 TTTTATCATTAAGGCTTCAAGATTGGCAAGTTCCAGCATGATATACAAATGCTTTTGGAA 2835
Qy 3576 AAATCAAGAGGAATTTCCCGAGTTAGAGGCCAGAGGACATGATTAACGTGGATCTTCAG 3635
Db 2836 AAATCAAGAGGAATTTCCCGAGTTAGAGGCCAGAGGACATGATTAACGTGGATCTTCAG 2895
Qy 3636 ATGTTTGACAGAGTGAAGGGATTAGAGAAAATCTTGTAAATTTGTTAGCAACCATCA 3695
Db 2896 ATGTTTGACAGAGTGAAGGGATTAGAGAAAATCTTGTAAATTTGTTAGCAACCATCA 2955
Qy 3696 GGATCGGAATCTTATTAAGAAATGATCAGATTAATCATGATTAAGAAAAGAGAGAGAGAAA 3755
Db 2956 GGATCGGAATCTTATTAAGAAATGATCAGATTAATCATGATTAAGAAAAGAGAGAGAGAAA 3015
Qy 3756 AGAAAAGCTGAAGCTGTAGGCTACATCGCAGAGAGATCATGGCTCAGATGTCTGCCCTTA 3815
Db 3016 AGAAAAGCTGAAGCTGTAGGCTACATCGCAGAGAGATCATGGCTCAGATGTCTGCCCTTA 3075
Qy 3816 CAGAAAACCTCATTTGAAACTCATTAACCTCATGTATGACATATACATACAGAAATGCTGGG 3875
Db 3076 CAGAAAACCTCATTTGAAACTCATTAACCTCATGTATGACATATACATACAGAAATGCTGGG 3135
Qy 3876 AAAGAAGATTCATTTATGAG 3935
Db 3136 AAAGAAGATTCATTTATGAG 3195
Qy 3936 GCTTTGGGCTCTTAAACGGGGTCCATCTGTCTGTAAGAGAGAGAGAGAGAGAGAGAGAGAG 3995
Db 3196 GCTTTGGGCTCTTAAACGGGGTCCATCTGTCTGTAAGAGAGAGAGAGAGAGAGAGAGAGAG 3255
Qy 3996 TGCCAAAGAGAACAGAGAGTGAATAATAGAAATATGCAATGATTTATCGGCTGTGTC 4055
Db 3256 TGCCAAAGAGAACAGAGAGTGAATAATAGAAATATGCAATGATTTATCGGCTGTGTC 3315
Qy 4056 CAGAAATCTATGCTTAAACCCAGCAGAGGGGAAACCCATAGAACTCTCAGAGAGAGAGCC 4115
Db 3316 CAGAAATCTATGCTTAAACCCAGCAGAGGGGAAACCCATAGAACTCTCAGAGAGAGAGCC 3375
Qy 4116 CTAGACCCACCTTTTTCATGGATCCAGATTTGGCATATGGAATTTATACAGGAGAGCTGTGGT 4175
Db 3376 CTAGACCCACCTTTTTCATGGATCCAGATTTGGCATATGGAATTTATACAGGAGAGCTGTGGT 3435
Qy 4176 CATGTAATGACAGCATGTGCTGGCAGAGATTTTGAAGCTGTACAGCTCAGCTTCAG 4235
Db 3436 CATGTAATGACAGCATGTGCTGGCAGAGATTTTGAAGCTGTACAGCTCAGCTTCAG 3495
Qy 4236 CAGCGCATTCATGTTGACCTTTTTCATGAGTGGAGAGATATCTTTGCGCTCTTTG 4295
Db 3496 CAGCGCATTCATGTTGACCTTTTTCATGAGTGGAGAGATATCTTTGCGCTCTTTG 3555

Qy 4296 AAATCTCTGTGCAATATCTGTGATCCCAATTAATCTTTTGCAACCTCAAAAGATAAACAGT 4355
Db 3556 AAATCTCTGTGCAATATCTGTGATCCCAATTAATCTTTTGCAACCTCAAAAGATAAACAGT 3615
Qy 4356 GAGAAATGAGATGCTCTGCTCAACTTTTGACCCCTGCGCACCGTGATACAGACTGCTTCG 4415
Db 3616 GAGAAATGAGATGCTCTGCTCAACTTTTGACCCCTGCGCACCGTGATACAGACTGCTTCG 3675
Qy 4416 GCCAGAAATATCAGGTTTATAATATAAGACATGCTAAAGGAGAAAACCCAAATTCCTATTTTC 4475
Db 3676 GCCAGAAATATCAGGTTTATAATATAAGACATGCTAAAGGAGAAAACCCAAATTCCTATTTTC 3735
Qy 4476 TTTAATCAAGGAATGAGGAGATTTCTACTTTTGAGATTCATTCATCTCAGATTTTGGCGTT 4535
Db 3736 TTTAATCAAGGAATGAGGAGATTTCTACTTTTGAGATTCATTCATCTCAGATTTTGGCGTT 3795
Qy 4536 GAGTCTTCGATTAATAATTTCAATAGCATCAAGGAAATGGTTATTTCTCTTTGGCCACACA 4595
Db 3796 GAGTCTTCGATTAATAATTTCAATAGCATCAAGGAAATGGTTATTTCTCTTTGGCCACACA 3855
Qy 4596 ATTTATAGAAATGGATTTGAAAGTCCACCTCATGAAAGGGATCTTCGAGTCCCAATGCTG 4655
Db 3856 ATTTATAGAAATGGATTTGAAAGTCCACCTCATGAAAGGGATCTTCGAGTCCCAATGCTG 3915
Qy 4656 ACCTGGAGCATCTGCGCTTTCATATCCAGGCAATTTGAAATCTATTTGGGAGATGAAGGA 4715
Db 3916 ACCTGGAGCATCTGCGCTTTCATATCCAGGCAATTTGAAATCTATTTGGGAGATGAAGGA 3975
Qy 4716 AAACCTCTGTTGGAGCACTTCAAAATAGGAGCATTAATGCTCTGAAAGCATTAATGAG 4775
Db 3976 AAACCTCTGTTGGAGCACTTCAAAATAGGAGCATTAATGCTCTGAAAGCATTAATGAG 4035
Qy 4776 TTTGAGTTGACAGAGGATTTACCTGCTCCTCAGGCTCTGATACAGAAACATCTGCTTCGT 4835
Db 4036 TTTGAGTTGACAGAGGATTTACCTGCTCCTCAGGCTCTGATACAGAAACATCTGCTTCGT 4095
Qy 4836 CTTCATCAGTTGTTCTTCTTAAACATAAAATCAGAAGATACACCATGCTTCTGCTATA 4895
Db 4096 CTTCATCAGTTGTTCTTCTTAAACATAAAATCAGAAGATACACCATGCTTCTGCTATA 4155
Qy 4896 GATCTGTTTCATGTTTGGTGGGCTGTGTAGCATTTCCCATCTCTGTATTTGGGATGAC 4955
Db 4156 GATCTGTTTCATGTTTGGTGGGCTGTGTAGCATTTCCCATCTCTGTATTTGGGATGAC 4215
Qy 4956 CCTGTTGATCTGACGCTTCTTCAAGTTAGTTCTTCTTAAACCATACCATCTTCTCTCCAT 5015
Db 4216 CCTGTTGATCTGACGCTTCTTCAAGTTAGTTCTTCTTAAACCATACCATCTTCTCTCCAT 4275
Qy 5016 TTGATCACCATGGCACAATGCTTTCAGATCTTACTTACAGTAGACACAGGCTTACCCCTT 5075
Db 4276 TTGATCACCATGGCACAATGCTTTCAGATCTTACTTACAGTAGACACAGGCTTACCCCTT 4335
Qy 5076 GCTCAGGTTTCAAGAGACAGTGAAGAGCTCATTTCCGATCTTCTTCTTTTGGCAGAAAT 5135
Db 4336 GCTCAGGTTTCAAGAGACAGTGAAGAGAGCTCATTTCCGATCTTCTTCTTTTGGCAGAAAT 4395
Qy 5136 TCTCAATATACAGTGGCTCCATTTGGTGTGATATTTCTGCTGTGATTTTGTGGTCTCA 5195
Db 4396 TCTCAATATACAGTGGCTCCATTTGGTGTGATATTTCTGCTGTGATTTTGTGGTCTCA 4455
Qy 5196 CTGAAGAAATGGCATCACCCCTTATCTTCCGCTGTGCTGATTTTTCCTACTATTTACTT 5255
Db 4456 CTGAAGAAATGGCATCACCCCTTATCTTCCGCTGTGCTGATTTTTCCTACTATTTACTT 4515
Qy 5256 GGGGTAACTCCGCTGAGGAACTGCATACCAATTTCTGCAAGAGAGAGATACAGTGCATC 5315
Db 4516 GGGGTAACTCCGCTGAGGAACTGCATACCAATTTCTGCAAGAGAGAGATACAGTGCATC 4575
Qy 5316 TGTAAGTATCTATCTTTACCTTACAAATTTGTTCTGCTCTTCCAGGATATTTGGGATCT 5375
Db 4576 TGTAAGTATCTATCTTTACCTTACAAATTTGTTCTGCTCTTCCAGGATATTTGGGATCT 4635

5376 GTAAAGCCCTTGGTCCAGAGGTGGTGCAGATCTCGCTTACTAAACCTGTTTGAAGCAA 5435
5377 |
5378 |
5379 |
5380 |
5381 |
5382 |
5383 |
5384 |
5385 |
5386 |
5387 |
5388 |
5389 |
5390 |
5391 |
5392 |
5393 |
5394 |
5395 |
5396 |
5397 |
5398 |
5399 |
5400 |
5401 |
5402 |
5403 |
5404 |
5405 |
5406 |
5407 |
5408 |
5409 |
5410 |
5411 |
5412 |
5413 |
5414 |
5415 |
5416 |
5417 |
5418 |
5419 |
5420 |
5421 |
5422 |
5423 |
5424 |
5425 |
5426 |
5427 |
5428 |
5429 |
5430 |
5431 |
5432 |
5433 |
5434 |
5435 |
5436 |
5437 |
5438 |
5439 |
5440 |
5441 |
5442 |
5443 |
5444 |
5445 |
5446 |
5447 |
5448 |
5449 |
5450 |
5451 |
5452 |
5453 |
5454 |
5455 |
5456 |
5457 |
5458 |
5459 |
5460 |
5461 |
5462 |
5463 |
5464 |
5465 |
5466 |
5467 |
5468 |
5469 |
5470 |
5471 |
5472 |
5473 |
5474 |
5475 |
5476 |
5477 |
5478 |
5479 |
5480 |
5481 |
5482 |
5483 |
5484 |
5485 |
5486 |
5487 |
5488 |
5489 |
5490 |
5491 |
5492 |
5493 |
5494 |
5495 |
5496 |
5497 |
5498 |
5499 |
5500 |
5501 |
5502 |
5503 |
5504 |
5505 |
5506 |
5507 |
5508 |
5509 |
5510 |
5511 |
5512 |
5513 |
5514 |
5515 |
5516 |
5517 |
5518 |
5519 |
5520 |
5521 |
5522 |
5523 |
5524 |
5525 |
5526 |
5527 |
5528 |
5529 |
5530 |
5531 |
5532 |
5533 |
5534 |
5535 |
5536 |
5537 |
5538 |
5539 |
5540 |
5541 |
5542 |
5543 |
5544 |
5545 |
5546 |
5547 |
5548 |
5549 |
5550 |
5551 |
5552 |
5553 |
5554 |
5555 |
5556 |
5557 |
5558 |
5559 |
5560 |
5561 |
5562 |
5563 |
5564 |
5565 |
5566 |
5567 |
5568 |
5569 |
5570 |
5571 |
5572 |
5573 |
5574 |
5575 |
5576 |
5577 |
5578 |
5579 |
5580 |
5581 |
5582 |
5583 |
5584 |
5585 |
5586 |
5587 |
5588 |
5589 |
5590 |
5591 |
5592 |
5593 |
5594 |
5595 |
5596 |
5597 |
5598 |
5599 |
5600 |
5601 |
5602 |
5603 |
5604 |
5605 |
5606 |
5607 |
5608 |
5609 |
5610 |
5611 |
5612 |
5613 |
5614 |
5615 |
5616 |
5617 |
5618 |
5619 |
5620 |
5621 |
5622 |
5623 |
5624 |
5625 |
5626 |
5627 |
5628 |
5629 |
5630 |
5631 |
5632 |
5633 |
5634 |
5635 |
5636 |
5637 |
5638 |
5639 |
5640 |
5641 |
5642 |
5643 |
5644 |
5645 |
5646 |
5647 |
5648 |
5649 |
5650 |
5651 |
5652 |
5653 |
5654 |
5655 |
5656 |
5657 |
5658 |
5659 |
5660 |
5661 |
5662 |
5663 |
5664 |
5665 |
5666 |
5667 |
5668 |
5669 |
5670 |
5671 |
5672 |
5673 |
5674 |
5675 |
5676 |
5677 |
5678 |
5679 |
5680 |
5681 |
5682 |
5683 |
5684 |
5685 |
5686 |
5687 |
5688 |
5689 |
5690 |
5691 |
5692 |
5693 |
5694 |
5695 |
5696 |
5697 |
5698 |
5699 |
5700 |
5701 |
5702 |
5703 |
5704 |
5705 |
5706 |
5707 |
5708 |
5709 |
5710 |
5711 |
5712 |
5713 |
5714 |
5715 |
5716 |
5717 |
5718 |
5719 |
5720 |
5721 |
5722 |
5723 |
5724 |
5725 |
5726 |
5727 |
5728 |
5729 |
5730 |
5731 |
5732 |
5733 |
5734 |
5735 |
5736 |
5737 |
5738 |
5739 |
5740 |
5741 |
5742 |
5743 |
5744 |
5745 |
5746 |
5747 |
5748 |
5749 |
5750 |
5751 |
5752 |
5753 |
5754 |
5755 |
5756 |
5757 |
5758 |
5759 |
5760 |
5761 |
5762 |
5763 |
5764 |
5765 |
5766 |
5767 |
5768 |
5769 |
5770 |
5771 |
5772 |
5773 |
5774 |
5775 |
5776 |
5777 |
5778 |
5779 |
5780 |
5781 |
5782 |
5783 |
5784 |
5785 |
5786 |
5787 |
5788 |
5789 |
5790 |
5791 |
5792 |
5793 |
5794 |
5795 |
5796 |
5797 |
5798 |
5799 |
5800 |
5801 |
5802 |
5803 |
5804 |
5805 |
5806 |
5807 |
5808 |
5809 |
5810 |
5811 |
5812 |
5813 |
5814 |
5815 |
5816 |
5817 |
5818 |
5819 |
5820 |
5821 |
5822 |
5823 |
5824 |
5825 |
5826 |
5827 |
5828 |
5829 |
5830 |
5831 |
5832 |
5833 |
5834 |
5835 |
5836 |
5837 |
5838 |
5839 |
5840 |
5841 |
5842 |
5843 |
5844 |
5845 |
5846 |
5847 |
5848 |
5849 |
5850 |
5851 |
5852 |
5853 |
5854 |
5855 |
5856 |
5857 |
5858 |
5859 |
5860 |
5861 |
5862 |
5863 |
5864 |
5865 |
5866 |
5867 |
5868 |
5869 |
5870 |
5871 |
5872 |
5873 |
5874 |
5875 |
5876 |
5877 |
5878 |
5879 |
5880 |
5881 |
5882 |
5883 |
5884 |
5885 |
5886 |
5887 |
5888 |
5889 |
5890 |
5891 |
5892 |
5893 |
5894 |
5895 |
5896 |
5897 |
5898 |
5899 |
5900 |
5901 |
5902 |
5903 |
5904 |
5905 |
5906 |
5907 |
5908 |
5909 |
5910 |
5911 |
5912 |
5913 |
5914 |
5915 |
5916 |
5917 |
5918 |
5919 |
5920 |
5921 |
5922 |
5923 |
5924 |
5925 |
5926 |
5927 |
5928 |
5929 |
5930 |
5931 |
5932 |
5933 |
5934 |
5935 |
5936 |
5937 |
5938 |
5939 |
5940 |
5941 |
5942 |
5943 |
5944 |
5945 |
5946 |
5947 |
5948 |
5949 |
5950 |
5951 |
5952 |
5953 |
5954 |
5955 |
5956 |
5957 |
5958 |
5959 |
5960 |
5961 |
5962 |
5963 |
5964 |
5965 |
5966 |
5967 |
5968 |
5969 |
5970 |
5971 |
5972 |
5973 |
5974 |
5975 |
5976 |
5977 |
5978 |
5979 |
5980 |
5981 |
5982 |
5983 |
5984 |
5985 |
5986 |
5987 |
5988 |
5989 |
5990 |
5991 |
5992 |
5993 |
5994 |
5995 |
5996 |
5997 |
5998 |
5999 |
6000 |

Query Match 48.4%; Score 3055.8; DB 17; Length 3059;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3057; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 680 GTCAAGTTCCTCAAGATGCGGACGAGGAGGCTGGAGGTACTGAGAGGATGGAATCAG 739
Db 1 GTCAAGTTCCTCAAGATGCGGACGAGGAGGCTGGAGGTACTGAGGAGGATGGAATCAG 60
Qy 740 GCGGAGTACCCGACGAGCCCTCAGCGTCTGGCATCTTGGTGGATCAGCAAGTTGATTT 799
Db 61 GCGGAGTACCCGACGAGCCCTCAGCGTCTGGCATCTTGGTGGATCAGCAAGTTGATTT 120
Qy 800 TTATCTGCTTCTTGGCATCAATTTGGCACAATTTGGTGGATGGAATTTTCTTGTGGAAT 859
Db 121 TTATCTGCTTCTTGGCATCAATTTGGCACAATTTGGTGGATGGAATTTTCTTGTGGAAT 180
Qy 860 GGACCCAGACTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 919
Db 181 GGACCCAGACTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240
Qy 920 ATGTGATCTTATTTGGAGAGATCCAGATATTTGGTGGATGGAATTTTCTTGTGGAAT 979
Db 241 ATGTGATCTTATTTGGAGAGATCCAGATATTTGGTGGATGGAATTTTCTTGTGGAAT 300
Qy 980 ATTTGAGCTTGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1039
Db 301 ATTTGAGCTTGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
Qy 1040 TGCAATTTGATCCAACTGATGATCTCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1099
Db 361 TGCAATTTGATCCAACTGATGATCTCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420
Qy 1100 TCATGTTTCAAGATGATCTTCTTCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1159
Db 421 TCATGTTTCAAGATGATCTTCTTCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480
Qy 1160 GGCAATGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1219
Db 481 GGCAATGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540
Qy 1220 AGAGAAATTCAGCTGCTCGGTTGAAATGAAGAGGATTAATTTGTCAGGAGGAGGAGGAGGAGG 1279
Db 541 AGAGAAATTCAGCTGCTCGGTTGAAATGAAGAGGATTAATTTGTCAGGAGGAGGAGGAGGAGGAGG 600
Qy 1280 TTCAATGATTAATTAATTTCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1339
Db 601 TTCAATGATTAATTAATTTCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
Qy 1340 TGAATCTCAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1399
Db 661 TGAATCTCAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720
Qy 1400 CCATTTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1459
Db 721 CCATTTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
Qy 1460 GGGGAGTTCATACCTACCTGCTTCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1519
Db 781 GGGGAGTTCATACCTACCTGCTTCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 840
Qy 1520 TTATGCTGCTTGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1579
Db 841 TTATGCTGCTTGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 900
Qy 1580 ACATCCACTTCATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1639
Db 1580 ACATCCACTTCATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG

Db 901 ACATCCACTTCATGTAGAGTATTACACTCAGAGATTATGGCTCATCAGAAATTTGCTTT 960
Qy 1640 GCGTCTTGCTGCTGGATGAACAAATATTAGAGTATTCAAGTGAATTTAGCAGATCTT 1699
Db 961 GCGTCTTGCTGCTGGATGAACAAATATTAGAGTATTCAAGTGAATTTAGCAGATCTT 1020
Qy 1700 TTGCCAAGCAGTCCCTTAGGAGAGACCTGACTCGGAGAAATCCCTGCTCATAAGCAGTT 1759
Db 1021 TTGCCAAGCAGTCCCTTAGGAGAGAACCTGACTCGGAGAAATCCCTGCTCATAAGCAGTT 1080
Qy 1760 AATGCTTTGGGATGCAAAAGCTTTATAAAGGTGCGGTAAAGATCTTTCAATGAATTCATCTT 1819
Db 1081 AATGCTTTGGGATGCAAAAGCTTTATAAAGGTGCGGTAAAGATCTTTCAATGAATTCATCTT 1140
Qy 1820 CAGCAGTCTTTTATGGAGATGGAATACAAAAACTCTTTGCTATGGAATTTGGAAGTA 1879
Db 1141 CAGCAGTCTTTTATGGAGATGGAATACAAAAACTCTTTGCTATGGAATTTGGAAGTA 1200
Qy 1880 TTATAAACAACGACGAAGATATATCAGTGAATGATCATGACAGAAATCTCTATAAC 1939
Db 1201 TTATAAACAACGACGAAGATATATCAGTGAATGATCATGACAGAAATCTCTATAAC 1260
Qy 1940 TGCACCTTCAGTTCAGATGTTTACTGTTCTCTACTCTGCTCGACATCTTATTGAAGACA 1999
Db 1261 TGCACCTTCAGTTCAGATGTTTACTGTTCTCTACTCTGCTCGACATCTTATTGAAGACA 1320
Qy 2000 GAATGTTATCTCTGTCATTACTGAATCTCTGCTAGAAGTTTTTACCTGAGTACTTTGACAG 2059
Db 1321 GAATGTTATCTCTGTCATTACTGAATCTCTGCTAGAAGTTTTTACCTGAGTACTTTGACAG 1380
Qy 2060 GAACATTAATTCACCTTCAGGTTTATAGCCAGGACAAATTTGGGAAGATATATGCACT 2119
Db 1381 GAACATTAATTCACCTTCAGGTTTATAGCCAGGACAAATTTGGGAAGATATATGCACT 1440
Qy 2120 AATATGTCACCTTAAAGTATATCTGATCAGCAAAACCCACAAATATGCGACAGAAAGTTAAG 2179
Db 1441 AATATGTCACCTTAAAGTATATCTGATCAGCAAAACCCACAAATATGCGACAGAAAGTTAAG 1500
Qy 2180 AATGAGTCTCTGAGGTTTTTTCGATCTTTTGAAGATCTTACCTGTATGCGAGGGAAT 2239
Db 1501 AATGAGTCTCTGAGGTTTTTTCGATCTTTTGAAGATCTTACCTGTATGCGAGGGAAT 1560
Qy 2240 GGAAGAAATCCGAGACAGGTTGGGCAACACATTTGAAGTGAATCTGATGGGAGGCTGC 2299
Db 1561 GGAAGAAATCCGAGACAGGTTGGGCAACACATTTGAAGTGAATCTGATGGGAGGCTGC 1620
Qy 2300 CATTTGCTATACAGATGCAATTTGAAGATATTTTACTCATGTTTCCAAAGAGTGGTGTGTTG 2359
Db 1621 CATTTGCTATACAGATGCAATTTGAAGATATTTTACTCATGTTTCCAAAGAGTGGTGTGTTG 1680
Qy 2360 TGATGAAGAACTTTTACTTTGGCTTTATAAAGAAATGTCAAAAGCTGTGATGAGGTGCAG 2419
Db 1681 TGATGAAGAACTTTTACTTTGGCTTTATAAAGAAATGTCAAAAGCTGTGATGAGGTGCAG 1740
Qy 2420 TACAGTTTCTATCTAGTAGCAAGACAGTATGATCAATCTGTTGGAACATAGTTTGGAAAC 2479
Db 1741 TACAGTTTCTATCTAGTAGCAAGACAGTATGATCAATCTGTTGGAACATAGTTTGGAAAC 1800
Qy 2480 AAGTCTCTACAGATGATCTGAGGATCTTGTATGACATACATCTGCCACTCTCTAGACCCCT 2539
Db 1801 AAGTCTCTACAGATGATCTGAGGATCTTGTATGACATACATCTGCCACTCTCTAGACCCCT 1860
Qy 2540 TGCTGGTCTTCAATGTACGTTTAAAGCAGGCTGGGTGCTGTTTCAAGACTGACATGAATTTGT 2599
Db 1861 TGCTGGTCTTCAATGTACGTTTAAAGCAGGCTGGGTGCTGTTTCAAGACTGACATGAATTTGT 1920
Qy 2600 GTCTTTTGAAGACTTTCAAGTAGAGGTACTAGTGAATATCTTTTACGTTCTCTGTTGTT 2659
Db 1921 GTCTTTTGAAGACTTTCAAGTAGAGGTACTAGTGAATATCTTTTACGTTCTCTGTTGTT 1980
Qy 2660 GGTGGCCAGGTTGTTGCTGAGATGTGGCGAAGAAATGGACTGTCTCTATTATGACAGGT 2719
Db 1981 GGTGGCCAGGTTGTTGCTGAGATGTGGCGAAGAAATGGACTGTCTCTATTATGACAGGT 2040

Qy 2720 GTTTTATTACCAAGATGTTTAAAGTGCAGAGAAATAATGTATGATTAAGATATCATCATGCT 2779
Db 2041 GTTTTATTACCAAGATGTTTAAAGTGCAGAGAAATAATGTATGATTAAGATATCATCATGCT 2100
Qy 2780 TCAGATTTGGTGCATCTTTTAAATGGATCCCAATAAGTTCTTTGTTACTGTGTACTTTCAGAGTA 2839
Db 2101 TCAGATTTGGTGCATCTTTTAAATGGATCCCAATAAGTTCTTTGTTACTGTGTACTTTCAGAGTA 2160
Qy 2840 TGAATCTTGGCAGGCTTTTAAACAAGACCATATCTACAAAAGACCAAGGATTTGATTTAAACA 2899
Db 2161 TGAATCTTGGCAGGCTTTTAAACAAGACCATATCTACAAAAGACCAAGGATTTGATTTAAACA 2220
Qy 2900 ATATAATACACTAATAGAGAAATGCTTCAGGTCTCTCATCTATATTTGGGTGAGCGTTA 2959
Db 2221 ATATAATACACTAATAGAGAAATGCTTCAGGTCTCTCATCTATATTTGGGTGAGCGTTA 2280
Qy 2960 TGTACTTGGAGTGGGAATGTGACCAAGAGAGGTCACAATGAGAGAAATCAATTCACCTT 3019
Db 2281 TGTACTTGGAGTGGGAATGTGACCAAGAGAGGTCACAATGAGAGAAATCAATTCACCTT 2340
Qy 3020 GCTTTGCAATGAAACCCATGCCACACAGTGCATTTGCCAAAATTTTACCTGAGAAATGAAAA 3079
Db 2341 GCTTTGCAATGAAACCCATGCCACACAGTGCATTTGCCAAAATTTTACCTGAGAAATGAAAA 2400
Qy 3080 TAATGAAACTGGCTTTAGAGAAATGTATATAAACAAGTGGCCACATTTAAGAAAACCAAGTGT 3139
Db 2401 TAATGAAACTGGCTTTAGAGAAATGTATATAAACAAGTGGCCACATTTAAGAAAACCAAGTGT 2460
Qy 3140 ATCAGGCCATGGAGTTTATGAACTTAAAGATGAAATCACTGAAAGCTTCAATATGTACTT 3199
Db 2461 ATCAGGCCATGGAGTTTATGAACTTAAAGATGAAATCACTGAAAGCTTCAATATGTACTT 2520
Qy 3200 TTATCATTTACTTCCAAAACCCAGCATAGCAAGCTTGAACATATGCAAGAAAGAGAGAAA 3259
Db 2521 TTATCATTTACTTCCAAAACCCAGCATAGCAAGCTTGAACATATGCAAGAAAGAGAGAAA 2580
Qy 3260 ACAAGAAAACAAGATGAAAGCATTTGCCGCCACCAACCTCTGAAATTTCTGCCCTGCTTT 3319
Db 2581 ACAAGAAAACAAGATGAAAGCATTTGCCGCCACCAACCTCTGAAATTTCTGCCCTGCTTT 2640
Qy 3320 CAGCAAGTGTAACTTCTCAACTGTGATATCATGATGTACATTTCTCAGGACCGTATTT 3379
Db 2641 CAGCAAGTGTAACTTCTCAACTGTGATATCATGATGTACATTTCTCAGGACCGTATTT 2700
Qy 3380 TGACGGGCAATAGACACAGATTTCTAACTTGTGACCGAAGGATGTCTCCAAATGGCTTT 3439
Db 2701 TGACGGGCAATAGACACAGATTTCTAACTTGTGACCGAAGGATGTCTCCAAATGGCTTT 2760
Qy 3440 TCATATTTCTGGCATTTGGTGTATAGAGAGAAAGCAACAGCTTCAAAAAGCTCTCTGAAGA 3499
Db 2761 TCATATTTCTGGCATTTGGTGTATAGAGAGAAAGCAACAGCTTCAAAAAGCTCTCTGAAGA 2820
Qy 3500 AGAAGTAACTTTGCACTTTTATCATAGGCTTCAAGATTTGGGAAGTTTTCAGCCATGAATAT 3559
Db 2821 AGAAGTAACTTTGCACTTTTATCATAGGCTTCAAGATTTGGGAAGTTTTCAGCCATGAATAT 2880
Qy 3560 ACAATCTTTTGGAAAACTCAAGGAAATTTCCCGAGTTTGAAGGCCAGAGGACATCAT 3619
Db 2881 ACAATCTTTTGGAAAACTCAAGGAAATTTCCCGAGTTTGAAGGCCAGAGGACATCAT 2940
Qy 3620 AACGTGGATACTTTCAGATTTTGAACAGTGAAGCGGATTTAAGAGAAAAAATCTTGTTTAAT 3679
Db 2941 AACGTGGATACTTTCAGATTTTGAACAGTGAAGCGGATTTAAGAGAAAAAATCTTGTTTAAT 3000
Qy 3680 TGTAGAACCAACATCAGATCGGATCTTATTAAGATGATGAGATTAATCATGATTAAG 3738
Db 3001 TGTAGAACCAACATCAGATCGGATCTTATTAAGATGATGAGATTAATCATGATTAAG 3059

RESULT 9
US-10-758-672A-3
; Sequence 3, Application US/10758672A

QY 2565 AGCTGGGTGCTGTTTCAAGACTGCATGAATTTGTGTCTTTTGGAGACTTTCAAGTAGAG 2624
Db 2173 AAAAGTGAAGTGGCATATATAAATTTCCAGAGCTCCTACCTCTAAGTGAACCTTAGCCACCC 2232
QY 2625 GTACTAGTGGAAATATCTTTACGTTGCTGTGGTGTGGTGGCCAGGTGTGCTGAGATG 2684
Db 2233 ATGTTGATAGAACACCTCTTAGAGTCTTGTCTGTGTGCCAGTACATGCGGAAATG 2292
QY 2685 TGGCGAAGAAATGGACTGTCTCTTAATAGCCAGGTGTTTATTAACAAGATGTTAAGTGC 2744
Db 2293 TGGAGAAGAAATGGGTCTCTCTAGTAAACCAAGATTTATTACTACCATTAATGTGAATGC 2352
QY 2745 AGAGAGAAATCATGATGAAGATATCATCATGCTTCAGATTTGAGTGTGCTATTAAGTAT 2804
Db 2353 AGACGTGAGATGTTTGACAAGGATGTAGTAATGCTTTCAGACGGTGTCTCCATGATGGAT 2412
QY 2805 CCCAATAAGTCTCTGTTACTGCTACTTCAGAGGTATGAATTTGCCAGGCTTTTAAACAG 2864
Db 2413 CCAAATCATTTCTGATGATCATGCTCAGCCGCTTTGAACCTTTATCAGATTTTCAGTACT 2472
QY 2865 ACCATATCTCAAAA-----GACCAGGATTTGATTTAAACAATAT 2903
Db 2473 CCAGACTATGAAAAGATTTAGTTCTGAGATTTACCCATAAGGATGTTGTTCCAGCAGAAC 2532
QY 2904 AATACACTAATAGAAGAAATGCTTCAGGCTCATCTATATTTGTGGGTGAGCGTTATGTA 2963
Db 2533 AATACCTCTAATAGAAAGATGCTATACCTCATTAATGCTTTGTTGGAGAGATTTAGT 2592
QY 2964 CCTGAGTGGGAAATGTGACCAAGAAAGAGGTGCACAAATGAGAGAAATCAATCACTTGCTT 3023
Db 2593 CCTGAGTGGACAGTAAATGCTACAGATGAATTAAGGATTAATCAAGGAGATTTCAATCAGTTG 2652
QY 3024 TGCAATTGAACCCATGCCACACAGTGCATTTGCCAAAATTTTACCTGAGAAATGAAAATAT 3083
Db 2653 AGTATCAAGCCTATGCTCATAGTGAATTTGGTAAAGTCTTTACCTGGAAGATGAGAACAAG 2712
QY 3084 GAACTGGCTTAGAGAAATGTCTATAAACAAGTGGCCACATTTAAGAACACAGGTGTATCA 3143
Db 2713 GAGACTGGCATGAGAGTGTATCGAAGCAGTGTGCCATTTTCAAGAAACCTGGAATTAACA 2772
QY 3144 GSCCATGGAGTTTATGAATGAATGAATCACTCACTGAAAGACTTCAATATGTACTTTAT 3203
Db 2773 GACAGGCAATGATGAATGAACTGAAACCAAGATTTGCCAAGATTTCAACTGTATTTCTAT 2832
QY 3204 CATTAATCCAAACCCAGCATAGCAAGGCTGAACATATGCAAGAAAGAGAGAAACAA 3263
Db 2833 CACTTTTCAAGGCGACAGCTCCAGGCAAGAGAGCGCAACGGAATTTGAAGACAA 2892
QY 3264 GAAACAAAGATGAAGCATTTGCCGCCACCAACCTCTCTGAATTTCTGCCCTGCTTTTACG 3323
Db 2893 AATAGAGAAGATACAGCACTCCCACTCCGGGTGTTGCCCTCCATTTCTGCCCTCTGTTTGA 2952
QY 3324 AAGGTGATTAACCTTCTCACTGATATCATGATGATCACTTCTCAGACCGTATTTGAG 3383
Db 2953 AGCTGGTTAAATTTGACGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 3012
QY 3384 CGGGCAATAGACACAGATTTCACTTTGGAACCGGAGGATGCTCCAAATGGCTTTTCTAT 3443
Db 3013 TGGGCTGTGACATATAATGATATGCTGCTGAGTCCATGCTGCAAGGAGTGTATCAT 3072
QY 3444 ATCTGGCAATGGGTTTACTAGAGAGAGCAACAGCTTTCAAAAAGCTCTCTGAAG---AA 3500
Db 3073 TTAATGGCATGGCACTCAAGAGAAAGAAACAAATTTAGAGATGTCAGCGAAGCAT 3132
QY 3501 GAAGTAAACATTTGATTTTATCATTAAGGCTTCAAGATTTGGAGATTTAGCCATGAAT--- 3557
Db 3133 GTAGTAAACATTTTACCTTCACTCAGAGATATCAAAAACCTGGTGAAGCGCCAAATAATCT 3192
QY 3558 -----ATACAAATGCTTTTGAAGAACTCAAGGAAATTTCCCAAGTGTAGAGGCCAGAG 3611
Db 3193 CCTACATACATGCTATGCTGGAACACATCAAAATGCTCTCCCTACTAGAGTCCACAAA 3252
QY 3612 GACATGATAACGTGGATATCTCAGATGTTTGAACACAGTGAAGCGGATTAAGAGAAATAATCT 3671

Db 3253 GACATGATTCGGTGGATATTTGAAGACTTTTAAATGCTGTTAAAGATGAGGAGATTTCA 3312
QY 3672 TGTTTAAATTTAGCAACCAACATCAGGATCGGAATCTTAAAGATGATGAGATTAATCTCAT 3731
Db 3313 CCTACCACTC-----CCGTGGCAGACAGAGGAAACCAATTAATGGAAGAGATTTCAAGG 3366
QY 3732 GATAAGAAAAGACAGAACGAAAAAGAAAGCTGGAAGCTGTAGGCTACATCCCGCAAG 3791
Db 3367 GACAAAGCAAGCTGTAGAGAGGAAAGAAAGCAGAGATTTGCCAGACTGCGCAGAGAAAAG 3426
QY 3792 ATCATGGCTCAGATGCTGCTCTTACAGAAAAAATTTCAATGAAATCATATAATCATATG 3851
Db 3427 ATCATGGCTCAGATGCTGCTGAAATGCGGGCATTTTATTTGATGAAAAACAAGACTCTTT 3486
QY 3852 GACATAATCATCAGAAATGCTGGGAAAGAAATTCATTTATGAGAAAGAGACACCCCA 3911
Db 3487 CAGCAGACATTAAGACTGATGCTCAACCTCTGCTGTCTTGTATCA-----TAGCCCT 3540
QY 3912 CGAGTCACTACTCTTAGAATTTGCTTGGGTCTTAAACGGGGTCCATCTCTTTACTGAA 3971
Db 3541 GTGGCTTCAGATATGACACTTACAGCACTGGGTCCACACAACTCAGGTTCTCTGAACAA 3600
QY 3972 AAGAGGTGCTGACGTGCATCTTTGCCAAGAAAGACAGGAGGTCAAAATAGAAAATAT 4031
Db 3601 AGACAAATTCGTTACATGTATTTGTCAAGAGGACGAAGATTTAAGTGGAAAGCAGG 3660
QY 4032 GCCATGTTATTCGGCTGTGTCAGAAATCTACTGCTTAAACCCAGCAGACGGGAAAA 4091
Db 3661 GCATGGTCTTGGCAGCATTTGTTCAAGATCACTGTATTAATCAAAAACAAGAGTAA 3720
QY 4092 CCCATAGAACTCTCAGGAGAACCCCTAGACCCATTTTCATGGATCCAGACTTGGCATAT 4151
Db 3721 TTTATTCAGATCCAGAAAA---TATGATCCATTTATTCATGCACCTGATCTGCTTGT 3777
QY 4152 GGAATCTTATCAGGAAGCTGTGCTCATGTAAATGACGAGGTGCTGCGCAGAGTATTTT 4211
Db 3778 GGAACACACATAGTAGTGTGGGCACATTTATGATGCCCATTTGTTGCCAAGAGTATTTT 3837
QY 4212 GAAGCTGTACAGCTGAGCTCTCAGCAGCGCATTCATGTTGACCTTTT-----T 4259
Db 3838 GATTCGGTTCAAGCTTAAGAACAGCGAAGGCAACAGAGATTAAGCTTACATACAGAGCTAT 3897
QY 4260 GACTTGGAAAGTGGAGAAATATCTTTGGCCCTCTTTGCAAAATCTCTGTCAATACTGTGATC 4319
Db 3898 GATGTAGAAAACGGAGAAATTCCTTTGCCCTTTTGTGAAATGCTTGAGTAACTACTGTATT 3957
QY 4320 CCCATATTTCTTTGCAACCTCAAAAGATAAAGTGAAGTGAAGTGCAGATGCTCTTCTCNA 4379
Db 3958 CC-----TCTGCTGCTCTCCAGAAATATTTTAAACAAGGTTAAATTTTTCAGAC 4011
QY 4380 CTTTTGACCTGGCAGCGGTGATACAGACTGTTCTGGCCAGAAATATCAGGTTTATAATATA 4439
Db 4012 CAACCAATCTGATCTCAGTGAATTAAGAACAAATATCTCAGCAANTTAAGCAATACAGTTT 4071
QY 4440 AGACATGCTAAGAGGAAACCCAAATTCCTATTTTCTTTAATCAAGGAATGGAGATTC 4499
Db 4072 CTTAGGAAGAGAAAGTACTCTCTAAATGCTCTCAAAAGATTCAGAAAAATGCGAT 4131
QY 4500 ACTTTGGAGTTCATTCATCTGAGTTTGGCGTTGAGTCTTCATTAATTAATTTCAAT 4559
Db 4132 GAAATACAGCTCCCTGAGGGTTTCAGGCTGATTTTTCGCTCTAAGATCCCTTATTTCTGAG 4191
QY 4560 AGCATCAAGGAAATGTTTATTTCTCTTGGCCACAAATTTATAGAAATTTGGAATTTGAAGTG 4619
Db 4192 AGCAATAAGAAATGCTAACGACATTTTGAACATGCTACTCAAGGTGGGATAAAGTTT 4251
QY 4620 CCACCTGATGAAGGGATTCCTCGAGTCCCATGCTGACCTGGAGCACCTGGGCTTTTCACT 4679
Db 4252 CATCCCAATGAAGAGATCTCTGTTCCCATTAATGTTGGGTAGCTGCGCGTACACC 4311
QY 4680 ATCCAGGCAATTTGAAATCTATTTGGAGATGAAGGAAACCTCTGTTTGGAGCACTTCA 4739

Db 4312 ATCCAAAGCATAGAAAGAAATTTGAGTGATGAAGATAAACCATTTGTTGGTCTTTTACCT 4371
Qy 4740 AATAGGACGATATATGGTCTGAAGCATTAATGCAAGTTTGGCAGTGGACAGAGATTACC 4799
Db 4372 TGCAGACTGGATGACTGTCTTAGGTGCAATGACGAGATTTGCCGACGACACTGGACAGTG 4431
Qy 4800 TGTCTCAGGTCTTGATACAGAAACATCTGGTTCGTCTTCTATCATAGTTGTTCTTCTTAAC 4859
Db 4432 GCATCAGTTTCAGTGGTGCAGGACATTTTGTAAACCTTTTGTGCATCACTGGTGCCTAAT 4491
Qy 4860 ATAAATCAGAAATACACCATGCTCTGCTATATAGATCTGTTTCATGTTTGGTGGGT 4919
Db 4492 GACAGCATGAGGAATCTTCATGCAATATAGATATTGACATGTTTCAATTTATTTGGTGGC 4551
Qy 4920 GCTGTGTAGCATTCCTCATCTGTTGATTTGGATGACCTGTTGATCTGCAGCCTTCTCA 4979
Db 4552 TTGGTGTGATTTCTGCGTGCAGTG-----TCAGGATTTTTCAGGG 4596
Qy 4980 GTTAGTCTCTCTATAACCATCTTTATCTCTTCCATTTGATCACCATGGCCACATGCTT 5039
Db 4597 ATCAGCCTTGGCAGTGGAGACCTTCACATTTTCCATCTGGTACTATGGGACACATCAT 4656
Qy 5040 CAGATPACTTACAGTACAGACAGCC- - -TACCCCTTGTCTCAGGTTCAAGAAGACAGT 5096
Db 4657 CAGATCTTACTTACCTCATGTACAGAGAAATGCGATGGATCAAGAAAATCCCCCTGT 4716
Qy 5097 GAAGAGGCTCATTCGGCATCTTCTTCTTGGAGAAATTTCTCAATATACAGTGGCTCC 5156
Db 4717 GAAGAAGAAATCAGCAGTCTTCTGTTGTATAAACAATCTTCCAGATATACGGGAAGTGCC 4776
Qy 5157 ATTGGGTGTGATATTCCTGCTGCTGATTTTGGGTCTCACTGAAAGATGGCATCACCCCT 5216
Db 4777 TTGAAGAAATACCATTCAGGCTGGCATCTGTGGAGAGTGTACAGCTGGAAATCATGCT 4836
Qy 5217 TATCTTCTGCTGTGCTGATTTGTTTCCATTTACTTTGGGTAACTCCGCTGAGGAA 5276
Db 4837 TTCTGAGTGTCTGCTTTATTTTCTTACTTTAAATGAGTTCTTCCACCCGAC 4896
Qy 5277 CTGATACCAATTTGCGAGAGGAGTACAGTGCATCTGTAGCTATCTATCTTTTACT 5336
Db 4897 ATTCA- - -AGTTCCTGGAAACAGCCATTTTGAACATTTTATGTAGCTATCTTCCCTACCA 4953
Qy 5337 ACATATTTGCTCTCTTCCAGCAATATTTGGGATCTGTAAGGCCCTTGTCCAGAGG 5396
Db 4954 AACACCTCATTTGCTTTTCAAGAAATAGTGAGATAATGAATTTCACTGATGAAGT 5013
Qy 5397 TGGTGTGAGATCTCTTAACTGTTTGAAGCAAAAAACACCGTGTGTCAGGTAC 5456
Db 5014 TGGTCCGTAAACAGTGAAGTTAAAGATATCTAGAAGGTGAAGAGATGCTATAAGATAT 5073
Qy 5457 CTTAGAAAAAGAAATAGTTGATAGAGCTTCTGTATGATATAGCTGCTCTGTGATCAA 5516
Db 5074 CCAAGAGAAATCTAAACAAATTAATAACCTTCCAGAGGATTAACAGAGCCTCATTAATCAA 5133
Qy 5517 GCTTCTCATTTTCAGGTGCCCAGGTCTGCAGATGATGAGCAAAACATCTGTCTCTGCG 5576
Db 5134 GCATCAAATTTCTGTGCCCCGAAATCAGTGGTGATTAAGACAGAGCCCCAACTCTGTGC 5193
Qy 5577 CTTTCTCTGTGGGCTATATCTATGTTCTCAGAAACATTTGCTGCGAGGAAATTTGTGAACGG 5636
Db 5194 CTTGTGTGGGATCTCTGTGCTCTCCAGAGTTACTGTGCCAGTGAATCTGGAAGGG 5253
Qy 5637 GAAGAGTTGGAGCTTGCAATTTTTCAGGACTTCTCATCTGTGAGCCGAGTCTGCATTTTC 5696
Db 5254 GAGGATGTAGGAGCCTGCACAGCTCACACCTACTCTCTGTGGCTCTGGAGTGGGCACTTTC 5313
Qy 5697 CTAAAAATCAGGAATGCGGTGCTCTGTTGAAGGTAAAGCCAGAGGCTGTGCCCTAT 5756
Db 5314 CTGAGAGTACGGGAATGTCAGGTGCTATTTTGTGCTGCAAAACCAAAGGCTGTTTTAT 5373
Qy 5757 CCAGCTCTCTTACTTGGATGAATATGGAGAAACAGACCTTGGCTGGAAGGGGCAACCCC 5816
Db 5374 TCTCCTCTTACTTGTGATCTATGGGAGACCGACCGGACTCAGACGGGAAATCCT 5433

Qy 5817 CTTCAATTTATCTGTGAGCGTATCGAAGCTCCATTTGTTGCTGGCAACACACTGCAAT 5876
Db 5434 TTACATTTATGCAAGAGCGAATTCAGAGATTCAGAGCTCTGGCACCAACACAGTGTG 5493
Qy 5877 ATAGAAGAGATTGCTAGGAGCCCAAGAGACTAATCAGATGTTTATTTGGATTCAACTGGCA 5935
Db 5494 ACAGAGAAATTTGACATGACAGAGGAAACCAATCAGACACTGGTTGGCACTTGACTGGCA 5552

RESULT 10

US-10-758-636A-3
; Sequence 3, Application US/10758636A
; Publication No. US20050089876A1
; GENERAL INFORMATION:
; APPLICANT: Han, et al.
; TITLE OF INVENTION: HUMAN E3 ALPHA UBIQUITIN LIGASE FAMILY
; FILE REFERENCE: 01017/35966C
; CURRENT APPLICATION NUMBER: US/10758,636A
; PRIOR FILING DATE: 2004-01-15
; PRIOR APPLICATION NUMBER: US 09/724,126
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 60/187,911
; PRIOR FILING DATE: 2000-03-08
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 6300
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (295)..(5559)
US-10-758-636A-3

Query Match 21.9%; Score 1381.8; DB 21; Length 6300;
Best Local Similarity 56.5%; Pred. No. 0;
Matches 2947; Conservative 0; Mismatches 2167; Indels 105; Gaps 16;

Qy 780 TGGATCAGCAGTTGATTTTATATCTGCTTTCTTGATCATTTTGGCAATTTGGTGCCA 839
Db 376 TGGTGAAGCACTGACCTCCTAGAGAGTGTACCAGCACTTTAGCCACTATGTATCCC 435
Qy 840 GAAATTTACTTTGCTGAAATGCGACCCAGACTTGGAAAGCAGGAGGAAGTGTACAAATG 899
Db 436 AAAATCTACTGCGGGTCCCAACCTTTTCCACAGAAAGAGACATGCTGGCACAGCAT 495
Qy 900 TCAATATTCACTCCACTGGAATGGTACTTATTTGGAGAAGATCCAGATATTTCCTTAGAG 959
Db 496 GTTTTGTGGACCAATGGAATGGTACTCTTTGTGTGAAGATCTTGCATTTTGGATTTTCCA 555
Qy 960 AAATTTGAAGCAC- - -AGTGGAGCAATTTTCAGCTTTTGGGAGGGTTTCAAAGTGGAGAG 1016
Db 556 AAATTTGAGCAAGCAAAACAACTCTCATCTTTGTGTGCTGTGTTTAAAGTAGGAGAG 615
Qy 1017 ACAACCTATTCTTCAGGGATTGTGCAATTCATCCAACTGTGTACTCTGTATGAGCTGC 1076
Db 616 CTTACATATTCTTGACAGACTGTGCACTTGTATCAACTGTGTGTTTGTGATGGAGTGC 675
Qy 1077 TTCCAGGACAGTGTTCATAAAATCATCGTTTACAAGATGCAATCTTCTACTGGAAGGGG 1136
Db 676 TTTTGGGAGATTTTACAGAGATCATCGATATAGGATGACAAACATCAGGAGGTGGAGGT 735
Qy 1137 TTCTGTGATCTGTGAGACACAGAGCATGGAATAATGCGCCCTTTTGTGTAAATCATGAA 1196
Db 736 TTCTGTGATCTGTGTGATCTACTGAAAGCCCTGGAAAGAGGGTCTTACTGTCAAACATGAA 795
Qy 1197 CCTGGAAGAGCAGGTACT- - -ATAAAGAGAAATTCAGCTGTCCGTTGGAATGAAGAGTGA 1253
Db 796 CTTAACCTCTGAAATTTGAGGAAGAGAGATCTCTTGTTCATTTATCAGAGATGTG 855
Qy 1254 ATGTCCAGCCAGGAAATATTTCTTCAAGTGATAAAATATGTGCTAGAAATGACTATA 1313

Db 856 ATAGCAAGAACTTATAACAATTTTCTATTAGTTTCGGTATGCGTAGAATAATTAAACC 915
Qy 1314 TGGGAAGAGGAAAGAAAGAACTGCTCTGAACTCCAGATAAGGGAGAGAAATAAGAAAGTAC 1373
Db 916 TGGGAAAGAAAGAAAGTAAATGCGAGCAGATTTAGAGATGGTAGAAGAGTGACACTTAC 975
Qy 1374 TATTGTGTCCTTTTCAATGATGAACACCAATTCATATGACCAGCTCATATACAGCTTCAA 1433
Db 976 TATTGCATGCTGTTTAAATGATGAGGTTTCAACCTATGAAACAGTATTTATATCTTTCAG 1035
Qy 1434 AGAGCTCTTGAGCTGCGAGCTGCGAGAGGCCAGTTGCGATACCACTGCGCAATGCAAAAGAG 1493
Db 1036 AAAGCTGTTAACTGTACACAAAAGAAAGCTATTGCTTTTGGCACTACAGTAGATCGAGAT 1095
Qy 1494 GGTGTCGGGCTGTTTAAAGCGGAGCTTATGCTGCTGCGAGGAAGCAAAAGAAAGATATA 1553
Db 1096 GGGCTAGGCTGTTGCGATATGAGATTTTCAAGTATGAGCAAGCAAAATCAGTAAT 1155
Qy 1554 AAGAGTCATTGAGAAATGTCTCTCAACATCCACTTCATGTAGAGTATTACACTCAGAG 1613
Db 1156 GTGAGAAATACAGTAGACAGAC---AAAGCCACTCAAAGTTTCAAGTTATGCAATTCGCT 1212
Qy 1614 ATTATGGCTCATCAGAAATTTGCTTGGCTTCTGCTGCTGATGAACAAAATTTATGAGC 1673
Db 1213 ATTGTGCGACATCAGAAATTTGGTTTGAACCTTTTGTCTTGGCTGGGAGTATTATTGA 1272
Qy 1674 TATTCAAGTGACTTTAGCGAGATCTTTTGGCAAGCATGCTTTAGAGAAAGAACTTGACTCG 1733
Db 1273 TATTCAAGATGCTTTCGCGCGATTTTATGTCAAAGTTGTTTACAAGAGGCCAGATGGT 1332
Qy 1734 GAGATCCCTGCTCATTAAGCAGGTTAATGCTTTGGGATGCAAGCTTTTAAAGGTGCC 1793
Db 1333 GAAAACTCTTCTAGTGGACAGACTGATGCTTAGTGATTCNAATTTAGGAAAGGTGCT 1392
Qy 1794 CGTAAGATCTTATGAATGATCTTACAGAGTTTATGAGAGTGAATACAAAAA 1853
Db 1393 AGGAGTGATATCATCAGTTGTTTATGAGCAGTCTGCTTATGATTTGAAATACAGAAA 1452
Qy 1854 CTCTTTGCTATGGAATTTGGAAGTATTATAAACAACCTGCAAGAAAGATATATCAGTGAT 1913
Db 1453 CTATTGCTGTTGATTTTGAAGAAATTTACAGCAGTTTGCAGAGAGATTTTATGGAGAT 1512
Qy 1914 GATCATGACAGAGTATCTTATTAACCTGCACTTTGAGTTTCAAGTTTCTGCTTACT 1973
Db 1513 GATCAGGAGCAGAGTGTGCGTGACTCTCTATCTGTCAGTTTCTTCAACGCGACTACT 1572
Qy 1974 CTGCTGACATCTTATTGAAGACAGAAATGTTATCTCTGTCATTCTGAACTCTGCTA 2033
Db 1573 CTGGCTGATGCTCATCACAGAGAAACCTTGATGAGCATTTATCAATTAAGACTTTTATG 1632
Qy 2034 GAACTTTT-----ACCTGAGTACTTGGACAGGAACAATAAATTCACCTCCAGGGTTAT 2087
Db 1633 GATCATTTGACACATCGAGATGCCAGGGCAGATTTCAAGTTTGAACGATACACTGCTTTA 1692
Qy 2088 AGCCAGGACAAATTTGGGAAGAGTATATGCGATATATATGATGACCTTAAGATATCTGATC 2147
Db 1693 CAAGCTTTCAAATTTAGGAGAGTACAGAGCCTTATTTTAGATCTCAAGTATGTTGTAAT 1752
Qy 2148 AGCAAAACCAATATGACAGAAAGATTAAGATGAGTTTCTTGAAGTTTTCGATCT 2207
Db 1753 AGCAAAACCACTGAATGTCAGATGAGCTGAGGAGAGATTTCTTAGAAGGTTTGTATGCC 1812
Qy 2208 TTTTGAAGATTTCTTACTGATGAGGAAATGGAAGAAATCCGAAGACAGGTTGGGCAA 2267
Db 1813 TTTTGGAAATTTACTAAATGATGCGAGGATGATCCAAATTAACGTCAGTAGGACAA 1872
Qy 2268 CACATTAAGTGAATCTGATTTGGAGGCTGCCATTTGCTATACAGATGCAATTAAGAAAT 2327
Db 1873 CATATTGAATGGAACCAAGAGTGGAGAGCCTTTCACACTACAAATGAATTAACACAT 1932
Qy 2328 ATTTTACTCATGTTCCAGAGTGTGCTTGTGATGAGAACTCTTACTTTGTGCTTAT 2387
Db 1933 GTCAATTTCAATGATGCGAGGACTGCTGCTTTCAGATGAAAGTGTTAATCGAAGCTTAC 1992

Qy 2388 AAAGAAATGTCACAAAGCTGTGATGAGGTGCGAGTACCACTTTTCATATCTAGTAGCAAGACA 2447
Db 1993 AAGAAATGCTCGCTGCTACTGATGATGAGTGTCAATGCTGTATATCTGATGCTGGAACAGCCA 2052
Qy 2448 GT---AGTACAAATCGTGTGGACATAGTTTGGAAACAAAGTCTCTACAGAGTATCTGAGAT 2504
Db 2053 ATCACACTAAGCAATTTTGGGACATTCAGTGGAAACTATCAGATATCTGTGTTTCCCAAGAA 2112
Qy 2505 CTTGTAAGCATACATCTGCCACTCTCTAGGACCCCTTGTGCTTCTCATGTAGCTTTAAGC 2564
Db 2113 AAAGTTAGCAATTCACCTCCAGTTTCTGCTTACTTTGAGGTTTACATGATATTTAAGC 2172
Qy 2565 AGGCTGGGTGCTGTTTCAAGACTGCAATGAAATTTGTGCTTTTGGAGACTTTTCAAGTAGAG 2624
Db 2173 AAAAGTGAAGTGGCATATAAATTTCCAGAGCTCTCACTCTAAGTGAACCTTAGCCCAACC 2232
Qy 2625 GTACTAGTGGAAATATCCTTTTACGTTGCTGTGTTGTTGCTGCGGAGTTTGTGCTGAGATG 2684
Db 2233 ATGTTGATAGAACACCTCTTTAGATGCTTGTCTGTGTCGCCAAGTACATGCGCGAATG 2292
Qy 2685 TGGGAAAGAAATGGAATGCTCTCTTATTAGCCAGGTTTATTATACCAAGATGTTAAGTGC 2744
Db 2293 TGGAGAGAAATGGTCTCTCTAGTAAACAGATTTATTACTACCATTAATGTGAAATGC 2352
Qy 2745 AGAGAAAGAAATGATGATGAATAAGATATCATCATGCTTCAAGATTTGTTGCAATCTTTAATGAT 2804
Db 2353 AGAGCTGAGATGTTTGAACAGGATGTAGTAATGCTTTCAGACAGGTGTCTCCATGATGAT 2412
Qy 2805 CCCAATAGTCTTGTACTTACTGTTTCAAGAGTATGAACCTTCCGAGGCTTTTAAACAG 2864
Db 2413 CCAATCATTTTCTGTGATGATCATGCTCAGCGCTTTTGAACCTTTATCAGATTTTCACTACT 2472
Qy 2865 ACCATATCTCAAAA-----GACCAGGATTTTGAATTAACAAATAT 2903
Db 2473 CCAGACTATGGAAGAAAGATTTAGTTCTGAGATTACCCATAGAGATTTGTTTACGCGAAGC 2532
Qy 2904 AATACATTAATPAGAAGAAATGCTTCAAGTCTCATCTATATTTGTTGGTGAGCGTTATGTA 2963
Db 2533 AATACTCTAATAGAAGAAATGCTATACCTCATTAATAATGCTTTGTTGGAGAGAGATTTAGT 2592
Qy 2964 CTTGAGTGGGAATGTGACCAAGAGAGGTGCACATGAGAGAAATCAATTCATCTGCTT 3023
Db 2593 CTTGGGTTGGACAGGTAATGCTACAGATGAATATCAAGCGAGAGATTTATCCATCAGTTG 2652
Qy 3024 TGCAATGCAACCAATGCGACACAGTGCCATTCGCAAAATTTTACCTGAGAAATGAATAAT 3083
Db 2653 AGTATCAAGCTATGGCTCATAGTGAATTTGTAAGTCTTTTACCTGAAGATGAGAACAG 2712
Qy 3084 GAAACTGGCTTATGAGAAATGTCATAAACAAGTGGCCACATTTTAAAGAAACCAAGGTGTATCA 3143
Db 2713 GAGACTGGCATGGAGAGTGTAAATCGAAGCAGTTGCCCATTTTCAAGAAACCTGGATTAACA 2772
Qy 3144 GGCCATGGAGTTTATGAATTAAGAGATGAATCATCTGAAAGACTTCAATATGCTACTTTTAT 3203
Db 2773 GGACGAGGCTATGATGAACCTGAAACCGAATGTGCCAAAGAGTTTCACTGTGATTTCTTAT 2832
Qy 3204 CATTACTTCAAAACCCAGCATAGCAAGCTAGCAAGCTGAACATATGCAAGAAAGAGAGAAACAA 3263
Db 2833 CACTTTTCAAGGGCAGAACAGTCCAGGCGAGAGAGGCGAACGGAAATTTGAAAGACAA 2892
Qy 3264 GAAACAAAGATGAAGCAATTTGCCCCACCACTCTCTGAAATTTGCTGCTTCTGCTTTCAGC 3323
Db 2893 AATAGAGAGATACAGCACTCCCACTCCGCTGTTGCTCTCCATTTCTGCTTCTGTTGCA 2952
Qy 3324 AAAGTGAATTAACCTTCTCAACTGTGATATCATGATGATACATTTCTCAGGACCGTATTTGAG 3383
Db 2953 AGCTGTTTAACTTTTGGAGTCAAGTGTATGTTGTCATCATGTTGGGAACTTTCTGCAA 3012
Qy 3384 CGGCAATAGACACAGATTTCTAATCTTGTGGACCGAAGGAGTGTCCAAATGCTTTTTCAT 3443
Db 3013 TGGGCTGTGGAACATAATAGGATATGCTTGTGTCAGATGCTGCTGCAAGAGGTGTATCAT 3072

Db 5194 CTTGTGTGGGATCTCTGCTGTGCTCCAGAGTTACTGCTGCCAGCTGAACCTGAAGGG 5253
Qy 5637 GAAGAGTTGGAGCTTTCATTTTTCACGCACTTCTCACTGTGGAGCCGAGTCTGCAATTTTC 5696
Db 5254 GAGGATGTAGGAGCTTCACAGCTCACACCTACTCTCTGTGGCTCTGGAGTGGGCACTTTC 5313
Qy 5697 CTAAATATCAGAGATGCGAGTGTCTCTGTGTTGAAGGTAAAGCCAGAGGCTGTCCCTAT 5756
Db 5314 CTGAGATACCGGAATGTCAAGTGTCTATTTTGTAGTGGCAAAACCAAGGCTGTTTTAT 5373
Qy 5757 CCAGCTCTTACTTCTGATGAATATATGAGAAACAGACCTTGGCTCAAGAGGGGCAACCC 5816
Db 5374 TCTCTCTTACTTGTATGACTATGTGGAGACCGACGAGTCTCAAGCGGGAATCT 5433
Qy 5817 CTTCAATTTATCTGTGAGCGGTATCGGAAGCTCCATTTTGGTCTGCAACCAACACTGCAAT 5876
Db 5434 TTACATTTATGCAAGAGCGATTCAAGAGATTCAAGAGCTCTGGCAACCAACAGATGTC 5493
Qy 5877 ATAGAGAGATGCTAGGAGCCAGAGACTAATCAGATGTTATTTGGATTCAACTGGCA 5935
Db 5494 ACAGAGGAAATTTGGACATGCAAGAGGCAATCAGACACTGGTTGGCAATTTGACTGGCA 5552

RESULT 11

US-10-758-672A-5

; Sequence 5, Application US/10758672A

; Publication No. US20040185037A1

; GENERAL INFORMATION:

; APPLICANT: Han, et al.

; TITLE OF INVENTION: HUMAN E3 ALPHA UBIQUITIN LIGASE FAMILY

; FILE REFERENCE: 010177/35966B

; CURRENT APPLICATION NUMBER: US/10/758,672A

; CURRENT FILING DATE: 2004-01-15

; PRIOR APPLICATION NUMBER: US 09/724,126

; PRIOR FILING DATE: 2000-11-28

; PRIOR APPLICATION NUMBER: US 60/187,911

; PRIOR FILING DATE: 2000-03-08

; NUMBER OF SEQ ID NOS: 29

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 5

; LENGTH: 6089

; TYPE: DNA

; ORGANISM: Mus musculus

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (766)..(6030)

US-10-758-672A-5

Query Match 19.6%; Score 1233.6; DB 19; Length 6089;

Best Local Similarity 54.7%; Pred. No. 0;

Matches 2870; Conservative 0; Mismatches 2269; Indels 105; Gaps 17;

Qy 769 TGGCATCTTGTGGATCAGCAAGTTGATTTTATCTGCTTCTTGATCATTTGGCAC 828
Db 836 TGCAGGAGATGGCTGCAAGCAACCGACCTCAACAGAGAGTGTACAGCAATTTAGCCC 895
Qy 829 AATTGGTCCAGAAATTTACTTTGCTGAAATGGACCCAGACTTGGAAAGCAGAGGAAA 888
Db 896 ACTGTGTCCCAAAATCTACTGCGGGGGCCCTTAACCCCTTCCCTCAGAAAGAGACACGC 955
Qy 889 GTGTACAAATGTCAATATTTCACTCCACTGGAAATGGTACTTATTTGGAGAAGATCCAGATA 948
Db 956 TGGCAGACACATCTCTGTGGACCGATGGAGTGTATCATCTGCGCTGAAGACCTGCGC 1015
Qy 949 TTTGCTTAGAGAAATTTGAAGCAC---AGTGAGCAATTTGAGTTTGTGGAGGGTTTCA 1005
Db 1016 TGGGATTTTCCAAAGCTCGAGCAGGCAAAACAAGCCCTTCTCACCTCTGTGGCCGAGTGTTA 1075
Qy 1006 AAGGTGAGAGACCACTTATCTTTCAGGAGATTTGCAATTTGATCCACATGTGTACTCT 1065
Db 1076 AAGTGGGGAAACCTACATCTCTCGAGAGCTGTGCAAGTGTGACCCCACTGTGTTTTAT 1135
Qy 1066 GTATGGACTGCTTCCAGGACAGTGTTCATAAAATCATCGTTACAGATGCATCTTCTA 1125

Db 1136 GCATGGAGTGTCTCTCTGGAAAGTATCCATAGAGACCATCGATATAGGATGACCACATCGG 1195
Qy 1126 CTGAGAGAGGGTCTGTGACTGTGGAGACACAGAGGCACTGGAAGACTGGCCCTTTTGTG 1185
Db 1196 GAGGAGGGGCTTCTGTGACTGTGGTGAACCTGAGGCGCTGGAAGAGGACCTTACTGCC 1255
Qy 1186 TAAATCATGAACCTGGGAAG---AGCAGGTACTATAAAAGAGAAATTCACGCTGTCCGTGA 1242
Db 1256 AGAAGCACAAGCTCAGCAGCTCTGAAGTTGTGGAGGAGGAGTCTCTTGTGCATCTAT 1315
Qy 1243 ATGAGAGGTAAATGTCCCAAGCCAGGAAATATTTCTTTCAGTGTATAAAATATGTCGTAG 1302
Db 1316 CAGAGATGTGATCGCCAGAACTTAACAATTTTGTCTATTTATGTATTTTCGATATGCAATAG 1375
Qy 1303 AAATGACTATATGGAAGAGGAAAGAACTGCGCTCTCTGAACTCCAGATAAAGGGAGAAA 1362
Db 1376 ATATACTGACCTGGGAAAGAAAGTGAATTTGCCCTGAAGACTTGAAGTGGCAGAGAAGA 1435
Qy 1363 ATGAAGATATCTATTTGTGCTCTTTCAATGATGAACACCATTTATATGACACACGTCATAT 1422
Db 1436 GTGACACCTACTACTGCACTGCTGTTTAAATGATGAGTTTCAACCTTATGAGCAAGTCAAT 1495
Qy 1423 ACAGCTACAAAGAGCTTCTGACTGTGAGCTCGCAGAGGCCAGTTGTCATACCATGCCA 1482
Db 1496 ATACCTCTCAGAAAGCTGTGAACCTGTACACAGAGGAAGCCATTGGCTTTTGCNACTACAG 1555
Qy 1483 TTGACAAAGGGTCTGTCGGGCTGTAAAGCGGAGCTTATGCTGCTTTCAGGAGGAGCA 1542
Db 1556 TTGATCGAGTGGCGGTAGGCTGTCCGATATGAGATTTCCAGTACTGTGATCAAGCNA 1615
Qy 1543 AGGAAGATATAAGAGTCAATTCAGAAAATGCTCTCAACATCCACTTTCATGTAGAAGTAT 1602
Db 1616 AGACAGTCAATTTGTGAGGAAACACAGCAGACAGACCAG---CCGCTCAAAGTTTCAAGTTA 1672
Qy 1603 TACACTCAGAGATATATGCTCATCAGAAATTTGCTTTGCGTCTTGGTCTCTGGATGAACA 1662
Db 1673 TGCACCTCTCGTGGCTGCTCATCAGAAATTTGGTTTGAAGACTCTGCTGCTGGCTGGAA 1732
Qy 1663 AAATATGAGCTATTTCAAGTCACTTTAGGCAGATCTTTTGGCAGACTGCTTACAGAGAG 1722
Db 1733 GTGTTATTTGGATATCTCAGATGGCTTTCGAGAGATTTTGTGTCNAGTTGATTTACAGAG 1792
Qy 1723 AACCTGACTCGGAGATCCCTGTCTCATAGCAGAGTAAATGCTTTGGGATGCAAAAGCTTT 1782
Db 1793 GTCCAGATGGGAAAACTCTTCTCTGGTGCAGACAGCTGATGCTTAATGATTTCCAAATAT 1852
Qy 1783 ATAAAGTGGCCGTAAGATCTTCAATGAATTTGATCTTTCAGCAGCTTTTATTTATGGAGTGG 1842
Db 1853 GGAAGGGGCTTAGGATGTGTATCACCAAGTTTGTTCATGAGCAGCTGCTCATGGACCTCA 1912
Qy 1843 AATACAAAAACTCTTTGCTATGGAATTTTGAAGTATTTATAAACAACCTGCAGAGAAAGAT 1902
Db 1913 AGTATAGAAGCTGTTCGCGCTTCGATTTGCTAAAATTTACCGCAGTTGCAGAGGATTT 1972
Qy 1903 ATATCAGTGTATGATCATGACAGAAAGTATCTCTATAACTGCACTTTTCAGTTTCAGATGTTTA 1962
Db 1973 TTATGGAGATGATCAAGGCGGCGAGTGTGGTGAAGTGTCTGTCTGTCTCCAGTTCTTCA 2032
Qy 1963 CTGTTCTTACTCTGGCTGCAATCTTATTTGAAGAGCAGAAATGTTATCTCTGTCTACTCTG 2022
Db 2033 CCGCACCGACGCTGGCGCAATGCTCTCTCAAGAGAGAACTCTGATGACCGTGTATCATTA 2092
Qy 2023 AAACCTGCTAGAGTTT-----ACCTGAGTACTTGGCAGGAAACAATAAATTTCAACT 2076
Db 2093 AGGCTTTTATGAGCAATTTTGAACAACAGAGATGCCAGGCGAGATTCAGTTTGAACGCT 2152
Qy 2077 TCCAGGGTTATAGCCAGGACAAAATTTGGAAGAGTATATGCAAGTAAATATGTCACCTTAAAGT 2136
Db 2153 ACCTGCGCTTCCAGCCCTTCAAGTTTCAGGAGATCCAGAGCTCATCTTAGATCTCAAGT 2212
Qy 2137 ATATCTGATCAGCAAAACCCCAAAATATGGAAGAGATTTAAGAAATGAGATGCTTCTTGAAG 2196

Db 2213 ATGATTGATTAGCAAAACCAACGAGTGGTCAGATGAGCTGAGGCAGGAAGTTCTTTACAAG 2272
Qy GTTTTCGATCTTTTGGAGATTCTTACTGTATGTCAGGGATGGAAGAAATCCGAGAC 2256
Db 2273 GGTTCGATGCCCTTCTTGAAATTACTGAAATGATGATGAGGAATGGACCCGATCAGCGTC 2332
Qy AGTTGGGCAACACATCAAGTGGATCCTGATTTGGAGGCTGCCATTGCTATACAGATGC 2316
Db 2333 AGTGGGACGACATTTGATGGAGCCAGAGTGGGAAGCAGCTTCACACTGACAGATGA 2392
Qy AATTGAAGAATAATTTTACTATGTTTCCAAAGTGGTGTGTTGATGAAGAATCTTTAC 2376
Db 2393 AGCTGACACACGTCTCAATGGTGCAGGACTGGTGTCTCTGACGAAAGATGTTAA 2452
Qy TTGTTGGCTTATAAGAAATGTCACAAAGCTGTGATGAGGTGACAGTTCATATCTA 2436
Db 2453 TTGAAGCTTACAAGAAATGCCGTGGCTGTGTGACACAGTGTGATGGCGGATTTACTGATG 2512
Qy GTAGCAAGACAGT---AGTACAATCGTGTGGACATAGTTTGGAAACAAGTCTTACAGAG 2493
Db 2513 GTGAACAGCCATCACACTCAGTATTTGTGGACACTCGGTGGAAACCATCAGATCTGTG 2572
Qy TATCTGAGGATCTTGTAAAGCATACATCTGCCACTCTCTPAGACCCCTGCTGTCTTCATG 2553
Db 2573 TTTCCCAAGAAAAGTTAGCATTTCACTCCCAATTTCTCGCTTGTCTGCAGGTTTGCATG 2632
Qy TACGTTTAAAGCGGCTGGTGTCTTCAAGACTGCATGAATTTGTGTCTTTTGAAGACT 2613
Db 2633 TATTGTTAAGCAAAAGTGAAGTGGCATATAAATTTCCAGAGCTCCTACTCTTAAGTGAAC 2692
Qy TTCAAGTAGAGTACTAGTGGAAATCTTTTACGTGTGTGTGTGTGGTGTGGCCAGGTG 2673
Db 2693 TGAGCCACCCATGTTGATGAGAAACATCTCTTAGATGTCCTTATGTCCTCAAGTGC 2752
Qy TTGCTGAGATGTGGCGAAGAAATGAGTGTCTCTTTATPAGCCAGGTGTTTATTAACCAAG 2733
Db 2753 ATGCTGGGATGTGGAGAAAGAAATGCTCTCTCTAGTAAATCAGATCTATTACTACCAT 2812
Qy ATGTTAAGTGCAGAGAAAGAAATGATGATGAAGATATCATCTGCTTCAGATTGTTGTCAT 2793
Db 2813 ATGTGAATGCGAGCGGAGAGATGTTTCSAAGGACATAGTATGCTTCAGACAGGTGCT 2872
Qy CTTTAAATGATCCCAATTAAGTCTTTTACTGTGTACTTTCAGAGGTATGAATCTTGGCCAGG 2853
Db 2873 CAGATGAGCCCAACCACTTCTCTGATGATCATGCTCAGCGCTTGTGAATCTCTATCAGC 2932
Qy CTTTAAACAAGACCATATCTACAAG-----ACCAGGATTTGA 2892
Db 2933 TCTTCAGCAGCGCTGACTATGGGAAGAGATTTCAGTTCTGAGGTTACCCATAAGGAGCTCG 2992
Qy TTAAACAATATATACATATAGAAAGAAATGCTTCAGGTCTCATCTATATTGTGGTG 2952
Db 2993 TTCAGCAGAAACAACACTCTGATCGAAGAGATGCTCTACCTCATCATCATGCTTGTGGAG 3052
Qy AGCGTTATGCTACCTCGAGTGGGAATGTGACCAAGAGAGGTCAACATGAGAGAAATCA 3012
Db 3053 AAGATTTCAACCTGGGTGGACAGGTGGTGGCCACAGATGAATTAAGAGGGAGATTA 3112
Qy TTCACTTGTGTTGCAATGAAACCATGCAAGTGGTGGCCAGTGTGCAAAAAATTTTACCTGAGA 3072
Db 3113 TCCATCAGTTGAGCATCAAGCTATAGGCTACAGTGAGTGGTGAAGTCTCTGCTGAAG 3172
Qy ATGAAAATTAAGAACTGGCTTAGAGATGTGATTAACAAAGTGGCCACATTTAAGAAC 3132
Db 3173 ATGAAACAAGGAGACCGCATGAGAGCGTCAATCGAGTCCGTTGCAATTTCAAGAAAC 3232
Qy CAGGTGTATCAGGCCATGAGTTTATGAATTAAGATGAATCACTGAAAGACTTCAATA 3192
Db 3233 CTGGCTCACAGGGCAGGATGTTATGAGCTGAGCCAGAGTGTGCCAAGAGTTCAACC 3292
Qy TGTACTTTTATCTACTCCAAAACCCAGCATPAGCAGGCTGAAACATATGCAAGAGAAAA 3252
Db 3293 TGTATTTTATCATTTCTCCAGGGCAGAGCAGTCCAAAGGAGGAGGCTCAGCGGAAAT 3352

Qy 3253 GGAAAAACAAGAAAAACAAGATGAAGCATTTGCCGCCACCAACCACTCTCTGAAATTCGACC 3312
Db 3353 TGAAGAAGAAAATAAAGAGATACAGCATCTCCCTCTCCGGCTTTGGCCACCGTTCTGCC 3412
Qy 3313 CTGCTTTCAAGAAAGTGAATTAACCTTCTCACTGTGATATCATGATGTACATTTCTCAGGA 3372
Db 3413 CTTTGTTCGGAGTCTGGTTAAATCTTTGCACTGTGACGTGATGCTGTGTACATCATGAGAA 3472
Qy 3373 CCGTATTTGAGCGGCAATAGACACAGATTTCTAACTTTGTGGACCGAAGGATGCTCCAAA 3432
Db 3473 CGATCTCTGCAAGTGGGCTGTAGAGCATCACGGGTCTGCTGTGTGAGTGTCCATGTCTACAGA 3532
Qy 3433 TGGCTTTTCTATTTCTGGCATTTGGGTTTACTAGAAAGAGAACACAGCTTCAAAAAGCTC 3492
Db 3533 GGGTGTCTGATTTGATCGGATGSGCTCTCCAGGAAGAGAGACCACTTGGAGAACGCG 3592
Qy 3493 CTGAAGAAGAGT---AACATTTGACCTTTTATCATAAAGGCTTCAAGATTTGGGAAGTTGAG 3549
Db 3593 TGAAGGCGCAGTGCAGACCTTCACTTCAACAGAAAGATTTCAAGGCTGTGTATGACAC 3652
Qy 3550 -----CCATGAATATACAAATGCTTTTGGAAAACTCAAAGGAATTTCCCAAGTTAG 3600
Db 3653 CACATAACTCCCCAGCATCTAGTATGCTGGAGACCTTGCAGAACGCCCTCCCTCG 3712
Qy 3601 AAGCCAGAGAGCATGATTAACAGTGGATATCTTCAAGATGTTTGGACACAGTGAAGCGATTAA 3660
Db 3713 AAGCCCAAGAGCATGATCAGGTGGTGTAAAGATGTTTAAAGCAATTAAGAGATTA 3772
Qy 3661 GAGAAAAATCTTGTAAATTTGTAGCAACCATCAGGATCGGAATCTTATTAAGAAATGATG 3720
Db 3773 GAGAGTGTTCATCCAGCAGCCCTGTGGCC-----GAGCGGAGGAAACCAATAATGGAGG 3826
Qy 3721 AGATTACTATGATAAAGAAAAAGCAGAAACGAAAAAGAAAGCTGAAGCTGTAGGCTTAC 3780
Db 3827 AGAGCTCAAGGACAGAGGACAAAGCAGAGAGGAAAGAAAGCCGAGATGCCAGACTGC 3886
Qy 3781 ATCCCCAGAGATCATGCTCAGATGCTGTGCTTTACAGAAAAAATTTCAATGAAACTCAT 3840
Db 3887 GCCGGAGAGATCATGCCCCAGATGCTGAGATGCGAGCGCATTTCAATGACGAAACA 3946
Qy 3841 AACTCATGTATGACATATACATCAGAAATGCTGGGAAGAAAGATTTCCATTTATGAGGAAG 3900
Db 3947 AAGAGCTTCTCCAGCAGACCCCTAGAGCTGGACACTCTGCTCTGCTGCACTCTTGAC---- 4002
Qy 3901 AGAGCAACCCAGCAGTCAAGTCACTCTAGAAATGCTTTGGGTCTTAAACGGGGTCCAT 3960
Db 4003 --AGCAGCCCTCCCGTTTCAGACGCGCTCTTACAGCACTGGGCCCCAGACACAGCAGG 4060
Qy 3961 CTGTTTACTGAAAAAGAGTGTGACGTGCATCTTTGCAAGAAAGAACAGAGGTTGAAAA 4020
Db 4061 TCCCTGAACCGAGACATTTGTCTACCTGTATATTATGTCAGAGAGGACAGAGGTGACTG 4120
Qy 4021 TAGAAAAATAATGCAATGATTTATTCGGCTGTGTCCAGAAATCTACTGCTTAAACCCAGC 4080
Db 4121 TGGGAAGCAGGCGCATGCTTTCGACGCTTTGTCAGAGGTCAACGGTCTCTGTCAAAG 4180
Qy 4081 ACAGGGGAAACCCATAGAACTCTCAGAGAGCCCTAGACCCACTTTTCATGATCCAG 4140
Db 4181 ACAGGACGAAACCATCGCGACCCAGAAAAATAT---GATTCATTTATTCATGACCCCG 4237
Qy 4141 ACTTGGCATATGAACTTTATACAGGAAGTGTGTGTCATGTAATGCAAGCAGTGTGCTGGC 4200
Db 4238 ATCTGTCTTGTGGGACACACACTGCGAGCTGTGGGACGTTATGATGCCCCATTTGTTGCG 4297
Qy 4201 AGAAGTATTTTGAAGCTGTAC-----AGCTGAGCTCTCAGCAGCGCATTCATGTTG 4251
Db 4298 AAGGTATTTTGAATTCGTTCAAGCAAGGAGCAGCAGAGGAGCAGCGGCTGGCTGTGC 4357
Qy 4252 ACCTT---TTTGACTTGGAAAGTGGAGATATCTTTGCCCTCTTTGCAAAATCTCTGTGCA 4308
Db 4358 ACATAGTACGATGTAGAGAAATGGGAGTTTCTCTGCCCCCTCTGTGAGTGCCTTGAGCA 4417

QY 4309 ATACTGTGATCCCATTTATCTTGGCAACCTCAAAAGATATAACAGTGAGATGAGATG 4368
DB 4418 ACACGGTGTATCCCC-----CTGCTGCTTCTCCAGGAGCATCTCCAGCAGGAGTTAA 4471
QY 4369 CTCCTGTCTCAACTTTTGAACCTTGGCAGCGGTGATACAGACTGTCTGGCCAGAAATATCAG 4428
DB 4472 ATTTTTCAGACCAACCAAGATCTGGCAGAGTGAACGAGAGCAGTAAACAGCAGATTAAGG 4531
QY 4439 GTTATAATATAAGACATGCTAAGAGAGAAACCCAAATTCCTATTTTCTTTAATCAAGAA 4488
DB 4532 TGGTCCAGATGCTGAGGAGAAAGCAATGCTGCTGACACAGCTCTTCTTCAGAGGACACAG 4591
QY 4489 TGGGAGATTTCTATTTGGAGTTCCATTCATCTGAGTTTGGCGTTGAGTCTTCGATTA 4548
DB 4592 AAGCCATGAATATAATACCGATCCCCGAGGCTTCAGGCGCTGATTTTATCCTAGNAACC 4651
QY 4549 AATATTTCAAAATAGCATCAAGGAATGGTTATTTCTCTTTGGCCACAACAAATTTATAGAAATG 4608
DB 4652 CATATTTCTGATAGCATATAAGGAATGTTAAACGACATTTGGAACGCTGCTTACAGGTGG 4711
QY 4609 GATTGAAGTGCACCTGATGAAGAGGATCTCTGAGTCCCCATGCTGACCTGGAGCACCT 4668
DB 4712 GACTGAAGTTCATCTTAATGAAGTGACCCCGTGTGCCATCTGTGCTGGGGGACCT 4771
QY 4669 GCGCTTTCATCTACAGGCAATGAATCTATTGGGAGATCAAGAAACCTCTGTTG 4728
DB 4772 GTGCATACACCATCCAGAGCATAGAAAGAAATTTTGAGTGTAGAGAGAGCCCTGTTTTG 4831
QY 4729 GAGCACTTCAAAATAGGAGCATATGCTCTGAAGCAATTAATGAGCTTTGAGTTGAC 4788
DB 4832 GACCTCTGCTTGTAGACTGAGAGCTGTCTCAGTGTGTTAACCGTTTGCAGAGCAC 4891
QY 4789 AGAGGATTAACCTGCTCCTCAGTCTCTGATACAGAAACATCTGTTGCTCTTCTATCAGTTG 4848
DB 4892 ATTGCACAGTGGCGTTACTTCTGCTGTGTAGAGGACACTTCTGTAAACTCTTTGTCATCT 4951
QY 4849 TTCTTCTCAATAAATACAGAGATACACATGCTCTGTCTATAGATCTGTTTCATG 4908
DB 4952 TGGTGCCTAGTGACAGCTATGAAGACCTCCCGTGCATCTAGACATCGACATGTTTCACT 5011
QY 4909 TTTTGGTGGGTGCTGTGTGTAGCATTCCTGATCCCATCTGTTATTTGGATGACCTGTTGATGCG 4968
DB 5012 TGCTGGTGGGCTGTGTGTCTGCTTTCCAGCTCTGAGTG-----TCAGG 5056
QY 4969 AGCCTTCTTCAGTTAGTTCTTCTCTATAACCACTTTATCTCTTCATTTGATCACCATGG 5028
DB 5057 ATTTTTCAGAGAGCAGCTGGCCACTGGGACCTGCACATCTTCCACTTGGTTACCATGG 5116
QY 5029 CACATGCTTCAGATACATCTTACAGTAGACACAGGCGCTA---CCCTTGTCTCAGGTTTC 5085
DB 5117 CACATGCTTACAGATCTTACTTACCTCATGTACAGAGAGAAATGGCATGGATCAAGAGA 5176
QY 5086 AAGAGACAGTGAAGAGCTCATTCGGCATCTTCTTCTTTTGGCAGAAATTTCTCAATATA 5145
DB 5177 ATCCCACTGGGGAAGAAAGAACTGGCCATCTCTTTTGGCAAAAACATCTCACCAAGTATA 5236
QY 5146 CAAAGTGGCTCAATTTGGGTGTATATCTCTGGCTGTATTTTGGGTCTCACTGAAGAAATG 5205
DB 5237 CTGGAAGTGCCTTGAAGAGAGCCCTTCGGCTGCACTGTGGAGGAGCGTCCGGGCGG 5296
QY 5206 GCATCACCCCTTATCTTCGCTGTGTCATTTGTTTTCATTTTATTTGTTGGGTAATC 5265
DB 5297 CCATCATGCTTTCTCAAGTGTCTGCTTTGTTTTCACATATTTTAAATGGAGTCCCGG 5356
QY 5266 GCGCTGAGGAAGTCAATACCAATTTCTGAGAGGAGTACAGTGCATCTGTAGCTATC 5325
DB 5357 CCCCTCCAGACCTTCA---AGTTTCTGGAACAAGCCATTTTGAACACATTTATGTAACTACC 5413
QY 5326 TATCTTTACCTACAAATTTGTTCTGCTCTTCCAGGAATATTTGGGATACTGTAAAGCCCT 5385
DB 5414 TTTCCCTACCAACCACTCATTCACCTTTTTCAGAAACAGTGCATCATGACTCCC 5473
QY 5386 TGCTCCAGAGGTGGTGTGAGATCTGCTCTTACTTAACTGTTTGAAGCAAAAAACACCG 5445

DB 5474 TGATTTGAAGTTGGTGGCCAGAACAGTGAAGTTAAACCGGTATCTAATATCGCAGAGAGGAG 5533
QY 5446 TGGTCAGGTACCTCTAGAAAAAAGAAATAGTTTGTATAGAGCTTCTCTGATGACTATAGCTGCC 5505
DB 5534 CGATAAGCTACCCACAGGAGGACTAACAACTGATAGACTTCCAGAGGATTTACAGCAGCC 5593
QY 5506 TCCTGAATCAAGCTTCTCATTTTCAGGTGCCCAAGGCTCTGCGAGATGATGAGCGAAAGCATC 5565
DB 5594 TCATTAACCAAGCATCCAATTTTCTGTGCCCAATCAAGTGGCCGACAAAGAGCAGAGCTC 5653
QY 5566 CTGTCTCTGCTCTTCTGTGGGGGTATACATTTCTCAGAACATTTTGTGCTGCCAGGAAA 5625
DB 5654 CTACTCTGTGCTCTGTGTGGAGTCTCTCTCTCTCAGAGTTACTGTGCTGCAAGCTG 5713
QY 5626 TTGTGAACGGGGAAGAGGTTGGAGCTTGCAATTTTTCACGCATCTTCACTGTGTGAGCGGAG 5685
DB 5714 AGCTGGAGGTGAGGAGCGTGGAGCGCTGCACAGCACACACCTACTCTCTCGGCTCCGGGG 5773
QY 5686 TCTGATTTTCTTAAATCAAGAAATGCCAGATGCCAGGTGCTCTGTTGAAGGTAAAGCCAGAG 5745
DB 5774 CCGGCATCTCTCAGAGTGGCGGAATGTCAAGTGTCTATTTTGTAGCTGGCAAAACCAAG 5833
QY 5746 GCTGTGCTATCCAGCTCTTACTTGGATGAATATGGAGAACACAGACCTTGCCTGAAGA 5805
DB 5834 GATGTTTATTTCTCTCTCTTACCTTGACGACTATGGAGAGACCGACCGGACTCTCAGAC 5893
QY 5806 GGGGCAACCCCTTCATTTATCTCTGAGCGGTATCGGAAGCTCCATTTGCTGTGGCAAC 5865
DB 5894 GAGGAATCTTTTACATTTATGCCAAGAGCGGTTTCCGAAGATCCAGAGCTCTGGCAGC 5953
QY 5866 AACACTGCTATTAAGAGAGATGCTAGAGCCAGAGACTAATCAGATGTTATTTGGAT 5925
DB 5954 AGCATAGTATCAAGAGGAGATCGACACGCGCAGGAGGCTTAACCCAGACCTGTGTCGAA 6013
QY 5926 TCAACTGGCAGTTACTGTGAGCTC 5949
DB 6014 TTGACTGGCAGCATTTTATAATCGC 6037

RESULT 12

US-10-758-636A-5

; Sequence 5, Application US/10758636A

; Publication No. US20050089876A1

; GENERAL INFORMATION:

; APPLICANT: Han, et al.

; TITLE OF INVENTION: HUMAN E3 ALPHA UBIQUITIN LIGASE FAMILY

; FILE REFERENCE: 01017/35966C

; CURRENT APPLICATION NUMBER: US/10758.636A

; CURRENT FILING DATE: 2004-01-15

; PRIOR APPLICATION NUMBER: US 09/724,126

; PRIOR FILING DATE: 2000-11-28

; PRIOR APPLICATION NUMBER: US 60/187,911

; PRIOR FILING DATE: 2000-03-08

; NUMBER OF SEQ ID NOS: 29

; SOFTWARE: Patent in version 3.2

; SEQ ID NO 5

; LENGTH: 6089

; TYPE: DNA

; ORGANISM: Mus musculus

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (766)..(6030)

US-10-758-636A-5

Query Match

Best Local Similarity 19.6%; Score 1233.6; DB 21; Length 6089;

Matches 2870; Conservative 0; Mismatches 2269; Indels 105; Gaps 17;

QY 769 TGGCATCTTGGTGGGATCAGCAAGTTGATTTTATCTGCTTTCTTGTCATCATTTGGCAC 828
DB 836 TCGCAGGAGGATGGCTGCAAGCAACCGGACCTCAACAGAGAAAGTGATCCAGCATTTAGCCC 895

829 AATTGGTGCAGAAATTTACTTTGCTGAAATGGACCCAGACTTGGAAAGCAGGAGGAAA 888
896 ACTGTGTGCCAAATCTACTGCGGGGCCCTTAACCCCTTCCCTCAGAGGAGACACGC 955
899 GTGTACAAATGTCAATATTTCACTCCACTGGAAATGTACTTATTTTGGAGAAATCCAGATA 948
956 TGGCAGACACATCTCTGCTGGGACCGATGGAGTGGTACATCTGCGCTGAAGACCCCTGGC 1015
949 TTTGCTTAGAGAAATTTGAAGCAC---AGTGGAGCAATTTTCAGCTTTGTGGAGGGTTTCA 1005
1016 TGGGATTTCCAAAGCTCGAGCAGGCAAAACAAGCCCTTCTCACTCTGTGGCCGAGTGTTA 1075
1006 AAAGTGGAGAGCAACCTATTCTTCAGGGATTTGTCAATGATCCAACTGTGTACTCT 1065
1076 AAGTGGGGAACCTACATACTCTCCAGAGACTGTGCAAGTTGACCCACCTGTGTAT 1135
1066 GTATGGACTGCTTCCAGGACAGTGTTCATAAAATCATCGTTTACAAAGATGCATCTTCTA 1125
1136 GCATGGAGTGTCTCTGGGAGTATCCATAGAGACCATCGATATAGGATGACCAATCGG 1195
1126 CTGGAGGAGGGTCTGTGACTGTGGAGACACAGAGGCAATGGAACTGGCCCTTTTGTG 1185
1196 GAGGAGGGGGCTTCTGTGACTGTGTGACACTGAGGCGTGGAAAGAGGAGCCCTTATGCC 1255
1186 TAAATCATGAACCTCGAAG---AGCAGGTACTATAAAAGAGAAATTCACGCTGTCGTTGA 1242
1256 AGAACCAAGCTCAGAGCTCTGAAGTTGTGGAGGAGAGGATCCTCTTGTGCATCTAT 1315
1243 ATGAAGAGGTAAATTTCCAAAGCCAGGAAATATTTCTTCAGTGATGATAAATATGTCGTAG 1302
1316 CAGAGAGTGTATGCCAGAACTTACAAATTTTGTCTATTATTTTTCGATATGATGATAG 1375
1303 AAATGACTATATGGGAGAGGAGAAAGAACTCCCTCTGAACTCCAGATATAGGAGGAGAAA 1362
1376 ATATACTGACCTGGGAGAAAGAGTGAATTCCTCTGAAGACTTAGAAGTGGCAGAGAGA 1435
1363 ATGAAGATACTATTGTGTCCTTTTCAATGATGAACACCATTCATATGACACGTCATAT 1422
1436 GTGACACCTACTACTGCAATGCTGTGTTAATGATGAGGTTTCACTTATGAGCAAGTCATTT 1495
1423 ACAGCCTTACAAGAGCTCTTGACTGTGAGCTCGCAGAGGCCAGTTGATACCACTGCCA 1482
1496 ATACCCTTTCAGAAAGCTGTGAACCTGTACACAGAGGAGGCAATTTGGCTTTTGCACATCAG 1555
1483 TTGACAAAGAGGGTCTGCGGCTGTGTTAAGCGGAGCTTATGCTGCTGCCAGGAGCA 1542
1556 TTGATCGAGATGCGCGTAGGCTGTCCGATATGGAGATTTCCAGTACTGTGATCAAGCAA 1615
1543 AGGAAGATATAAGAGTCAATTCAGAAATGTCTCTCAACATCCACTTCATGTAGAAGTAT 1602
1616 AGCAGTCAATTTGGAGAAACACAGCAGAGAGACCAAG---CCGCTCAAGTTCAAGTTA 1672
1603 TACACTCAGAGATATATGGCTCATCAGAAATTTGCTTTGCGTCTTGCTTCTGGATGAACA 1662
1673 TGCACTCCCTCGTGGCTGCTCATCAGAAATTTGGTTTGAAGCTCTGCTGCTGGCTGGGA 1732
1663 AAATATGAGCTATTCAAGTGAATTTAGCAGATCTTTTGGCAAGCATGCTCTAGAGAG 1722
1733 GTGTTATTTGGATPACTCAGATGCGCTTCGAGGATTTTGTGTCAAGTTTGAATTTACAAAG 1792
1723 AACCTGACTCGAGAACTCCCTGCTCATAGCAGGTTAATGCTTTGGGATGCAAGCTTTT 1782
1793 GTCCAGATGGGCAAACTCTTCTCTGGTGCAGAGACTGATGCTTAAATGATTTCAAATAT 1852
1783 ATAAAGGTGCGGTGAAGATCTTTCATGAATTTGATCTTCAGCAGTTTTTTTTATGGAGATGG 1842
1853 GGAAGGGGCTAGGAGTGTATCACCAAGTTGTTTCAAGCAGCCTGCTCATGGACCTCA 1912
1843 AATACAAAAATCTTTGCTATGGAATTTGTGAAGTATTTATTAACACTGCGAGAAAGAT 1902
1913 AGTATAAGAAAGCTGTTGCGGCTTCGATTTTGTGTAATAAATTTACCGGAGTTGCGAGGGAT 1972
1903 ATATCAGTGTATCATGACAGAGTATCTCTATAACTGCACTTTTTCAGTTTCAGATGTTTA 1962

1973 TTATGGAGATGATCAAGAGCGGCAGTGTGCGTGACTGCTCTGCTCTGCTGCTCAGTTCTTCA 2032
1963 CTGTTCTTACTCTGCTGCGATCTTATTTAGAGAGCAGAAATGTTATCTCTGCTGCTACTG 2022
2033 CCGCACCAGCGCTGCGCGAATGCTCTCTCAGAGAGAGAACTCTGATGACCGCTTATCATTA 2092
2023 AAACTCTGCTAGAAAGTTTTT-----ACCTGAGTACTTTGGACAGGAAACAATAAATTTCAACT 2076
2093 AGGCTTTTCATGGACGATTTTGAACAACAGAGATGCCAGGCGAGATTCAGTTTGAACGCT 2152
2077 TCCAGGGTTATPAGCCAGGACAAAATTTGGGAAGAGTATATGCAAGTAAATATGTGACCTTAAAGT 2136
2153 ACATGCGCTTCAAGCTTCAAGTTTCAGGAGATCCAGAGCCTCATCTTAGATCTCAAGT 2212
2137 ATATCTGTATCAGCAAAACCCCAATATATGGAAGAGAAATTAAGAAATGAGTTTCTTGAAG 2196
2213 ATGTATTTGATTTAGCAAAACCAACGAGTGGTCCAGATGAGCTGAGGCGAGAAATTTTACAAAG 2272
2197 GTTTTCGATCTTTTTCGAAAGATTTCTTACCTGTATGCAAGGAAATGGAAGAAATTCGAAGAC 2256
2273 GGTTCGATGCTCTTCTTGAATTTACTGAGTGCATGCAAGGAAATGGAACCCGATCAGCGTC 2332
2257 AGTGTGGGCAACACATTTGAAGTGGATCCTGATTTGGAGGCTGCCATTTGCTATACAGATGC 2316
2333 AGTGGGACAGCACATTTGAGATGAGCCAGAGTGGGAAGCAGCCTTCACACTGCGAGATGA 2392
2317 AATTGAAGATATTTTACTCATGTTCAGAGTGGTGTGCTGTGATGAGAACTCTTAC 2376
2393 AGCTGACACAGCTCATCTCAATGTTGCGAGGACTGGTGTGCTCTGGACGAAAGGTGTTAA 2452
2377 TTGTGGCTTATAAGAAATGTCACAAAGCTGTGATGAGTGCAGTACCAGTTTTCATATCTA 2436
2453 TTGAAGCTTACAGAAATGCTGCTGCTGTGTCACAGTGTTCATGGCGGATTTTACTGATG 2512
2437 GTAGCAAGACAGT---AGTACAATCGTGTGACATAGTTTGGAAACAAAGTCTTACAGAG 2493
2513 GTGAACAGCCAACTCACACTCAGTATTTGTGACACTCGGTGGAACCATCAGATACTGTG 2572
2494 TATCTGAGGATCTTTGTAAGCATACATCTGCCACTCTCTAGGACCTTGTGCTGCTTCAATG 2553
2573 TTTCCCAAGAAAAAGTTAGCATTCACCTCCCAATTTCTCGCTTGTGTCAGGTTTGCATG 2632
2554 TACGTTTAAAGCAGGCTGGTGTCTGTTTCAAGACTGCATGAATTTGTGCTTTTGAGGACT 2613
2633 TATTGTTAAGCAAAAGTGAAGTGGCATATAAATTTCCAGAGCTCCTACCTCTAAGTGAAC 2692
2614 TTCAAGTAGAGGTACTAGTGGAAATATCCTTTTACCGTTGTCTGCTGTTGGTTGCCAGGTTG 2673
2693 TGAGCCCAACCATGTTGATAGAACATCCTCTTAGATGTCTTGTCTTATGTCTCAAGTGC 2752
2674 TTGCTGAGATGTGGGAGAAATGGAATGGAATGCTCTTATTTAGCCAGGTTTATTTACCAAG 2733
2753 ATGCTGGGATGTGGAGAGAAATGGCTTCTCTAGTAAATCAGATCTATTATACCATTA 2812
2734 ATGTTAAGTCAGAGAGAAATGATGATAAGATATCATCATGCTTCAGATTTGTGTCAT 2793
2813 ATGGAATGAGCGGAGAGATGTTGCAAGAGACATAGTATGCTTCAGACAGTGTCT 2872
2794 CTTTAAATGGATCCCAATAAGTTTCTTTTACTTGGTACTTTTTCAGAGGTATGAACTTGCAGG 2853
2873 CCATGATGGACCCAAACCACTTCTCTGATGATCATGCTCAGCGCTTTTGAACCTCTATCAGC 2932
2854 CTTTAAACAGACCATATCTACAAAG-----ACCAGGATTTGA 2892
2933 TCTTCAGCACGCTGACTATGGGAGAGATTCAGTTCTGAGGTTTACCCTAAGGACGTCG 2992
2893 TTAACATATAATACACTAATAGAGAAATGCTTTCAGGCTCCTCATCTATATTTGTGGTGT 2952
2993 TTCAGAGAACACACTCTGATCGAAGAGATGCTTCTACCTCATCATCATGCTTGTGGAG 3052
2953 AGCGTTATGTACCTGGAGTGGGAAATGTGACCAAGAGAGGTCACAATAGAGAGAAATCA 3012

Db 3053 AAGATTCAACCCCTGGGTTGGACAGGTGGCTGCCACAGATGAATCAAGAGGAGATTA 3112
Qy 3013 TTCACTGCTTTGCAATTGAACCCATGCGCCACACAGTGGCCATTGCCAAAATTTACTGTAGA 3072
Db 3113 TCCATCAGTGTAGCATCAAGCTATAGGCTCAAGTGAGCTGCTGAGTCTCTGCTCGAAG 3172
Qy 3073 ATGAAATTAATGAACTGGCTTAGAGATGTCTAAACAAAGTGGCCACATTTAGAAAC 3132
Db 3173 ATGAAACAAGGAGACCGGCTAGAGAGCGTCAATCGAGTCCGTTGCACATTTCAAGAAAC 3232
Qy 3133 CAGGTGTATCAGGCCATGAGTTATGAATCACTGAAAGATGAATCACTGAAAGCTTTCAATA 3192
Db 3233 CTGGCTCACAGGCGGAGCATGTATGAGTCAAGCCAGAGTGTGCCAAGAGTTCAACC 3292
Qy 3193 TGTATCTTTATCATTTCTCCAAAACCCAGCATAGCAAGCTGAAATATGCAAGAAAGAAA 3252
Db 3293 TGTATTTTATCATTTCTCCAGGCGAGAGCAGTCCAAAGCAGAGGAAGCTCAGCGGAAT 3352
Qy 3253 GGAGAAACAGAAACAAAGATGAAGCAATTCGCGCCACCACTCTGTAATTTCTGCC 3312
Db 3353 TGAAGAGAGAAATAAAGAGATACAGCACTCCCTCCCGGCTTTGCCACCGTTCTGCC 3412
Qy 3313 CTGCTTTCCAGCAAGATGAATTAACCTTTCTCAACTGTGATATCATGATGTACATTTCTCAGA 3372
Db 3413 CTTTGTTCGCGAGTCTGTTAACTCTTCAGTGTGACGTCAATGCTGTACATCATGGA 3472
Qy 3373 CCGTATTTGAGGGGCAATAGACACAGATTTCTAACTTGTGACCCGAAGGATGTCCAAA 3432
Db 3473 CGATCTGAGTGGGCTGTAGAGCATCAGGGTCTGCTGCTGAGTCCATGCTACAGA 3532
Qy 3433 TGGCTTTTATATCTGCGATTTGGGTTTACTAGAGAGAGAGCAAGCTTCAAAAGCTC 3492
Db 3533 GGGTGTCTGATTTGATCGGGATGGCTCTCCAGGAAGAGAGCAACCTTTGGAGAACGCGC 3592
Qy 3493 CTGAAGAGAGAGT---AATTTGACTTTTATCATTAAGCTTTCAAGATTTGGAGAGTTCCAG 3549
Db 3593 TGAAGGCGACGTGCAGACCTTCACTTCAACAGAGATTTCAAGCTGTGTATGAC 3652
Qy 3550 -----CCATGAATATACAAATGCTTTTGGAAAACTCAAGGAATTTCCCAAGTTAG 3600
Db 3653 CACATAACTCCCGAGCATCTAGCTATGCTGGAGACCTTGACAGAACGCCCTCCCTGG 3712
Qy 3601 AAGGCCAGAGCATGATTAACGTGGATCTTCAGATGTTTGCACAGTGAAGCATTA 3660
Db 3713 AAGCCACAGAGACATGATCAGGTGGTGTCTAAAGATGTTTAAATGCAATTAAGAGATA 3772
Qy 3661 GAGAAATCTTGTTTAAATGTAGCAACACATCAGGATCGAATCTTATTAAGATGATG 3720
Db 3773 GAGAGTTCATCCAGACGCTGTGGC-----GAGCGGAGGGAACCAATATGGAGG 3826
Qy 3721 AGATTACTCATGATAAGAAAGCAGAACGAAAGAAAAGCTGAAGCTGTAGGCTAC 3780
Db 3827 AGAGCTCAAGAGACAAGCAAGCAGAGAGAGGAAAGAAAGCCGAGATCGCCAGCTGC 3886
Qy 3781 ATGCGCAGAGATCATGCTGAGTGTCTGCTGCTTACAGAAATCTTCAATTTGAAATCTATA 3840
Db 3887 GCCGGAGAGAGATCATGCGCCAGAGTGTCTGAGATGACGGGCACTTCAATTCACGAAACA 3946
Qy 3841 AACTCATGTATGACATACATCAGAAATGCTGGGAAAGAGATTTCCATTTATGGAGAG 3900
Db 3947 AAGAGCTCTTCAGCAGACCTTAGAGCTGGACACCTCTGCTCTGCCACTCTTGAC----- 4002
Qy 3901 AGAGCAACCCAGCAGCTCAGTGAATCTTCTAGAAATTTGGGCTCTTAAAGGGGTCCAT 3960
Db 4003 ---AGCAGCCTCCCGTTTCAGAGCGAGCTCTTACAGCACTGGGCCAGCAGACAGCAGG 4060
Qy 3961 CTGTTTCTGAAAGAGAGTGTGAGTGTGATCTCTTTTGCAGAGAAACAGAGAGTGA 4020
Db 4061 TCCCTGAAACCGAGACAGTGTGTACCTGTATATATATGTTCAAGAGAGGCAAGAGGTGACTG 4120
Qy 4021 TAGAAATTAATGCCATGATTTATCGGCTGTGTCAGAAATCTACTGCTTAAACCCAGC 4080
Db 4121 TGGGAAGCAGGCGGATGCTTGGCAGCGTTTGTTCAGAGGTCAACGGTCTCTGTCAAAG 4180

Qy 4081 ACAGGGGAAAACCCATAGAACTCTCAGGAGAGCCCTAGACCCTTTTCTATGATTCAG 4140
Db 4181 ACAGAGCAAAAACCATCGCGACCCAGAAAAATAT---GATCCATTATTTATGCAACCCG 4237
Qy 4141 ACTTGGCATATGGAACTTATACAGGAAGCTGTGGTCATGTAATGCAAGGAGTGTCTGGC 4200
Db 4238 ATCTGTCTTGTGGGACACACACTGCGAGCTGTGGGACGTTATGATGCCCATTTGTCG 4297
Qy 4201 AGAAGTATTTTGAAGCTGTAC-----AGCTGAGCTCTCAGCAGCGCATTTCAATGTTG 4251
Db 4298 AAAGGTATTTTGAATTCGGTTCAAGCAAGGAGCAGGAAGCAGCAGCGCTCGCTTGC 4357
Qy 4252 ACCTT---TTTGAATTTGAAAGTGAAGATATCTTTGGCCCTCTTTGCAAAATCTCTGTGA 4308
Db 4358 ACATCTAGCTACGATGTAGGAATGCGAGTCTCTGCGCGCTCTGTGAGTGTCTTGAGCA 4417
Qy 4309 ATACTGTGATCCCATTTATCTTTTGAACCTCTCAAAAGATAAAGAGTGAAGTGCAGATG 4368
Db 4418 ACAGGTTATCCCC-----CTGCTGCTTCTCCAGGAGCATCTCAGAGGAGGTTAA 4471
Qy 4369 CTCTTGCTCAACTTTTGAACCTGCGCAGCGTGGATACAGACTGTTCTGCGCCAGAATATCAG 4428
Db 4472 ATTTTTCAGACCAACCCAGATCTGGCACAGTGGACGAGAGCAATACACAGCAGATAAAG 4531
Qy 4429 GTTATATATTAAGACATGCTAAAGAGAAACCCAAATTCCTATTTCTTTTAAATCAAGAA 4488
Db 4532 TGGTCCAGATCTGAGGAGAAAGCAATGCTGTGACACAGTCTTCTTTCAGAGGACACAG 4591
Qy 4489 TGGGAGATTTCTACTTTTGGAGTTCCATTCCTCAGTGTGTTGGGTTTGGGCTTCTGATTA 4548
Db 4592 AAGCCATGAATATTAATACCGATCCCGAAGGCTTCAGGCTGATTTTATCTTAGAAC 4651
Qy 4549 AATATTCAAAATAGCATCAAGGAATGTTATTTCTTTGGCCACAACTTTATAGAAATG 4608
Db 4652 CATATCTGATAGCATAAAGAAATGTTAAAGCATTTGGAAACGCTCTTACAGGTGG 4711
Qy 4609 GATTGAAAGTGGCACTGTATGAAGAGGATCTCGAGTCCCACATGCTGACCTTGGAGCACT 4668
Db 4712 GACTGAAGTGTCACTCTAAATGAAGTGAACCCCGTGTGCCCATCTCTGTCTGGGGGACCT 4771
Qy 4669 GCGCTTTCATCTCCAGCAATTCGAAATCTATTTGGGAGATGAAGGAATCTCTGTTTG 4728
Db 4772 GTGCATACACATCCAGAGCATAGAAAGATTTTGAAGTGAAGAGGAGGCTGTTTTTG 4831
Qy 4729 GAGCACTTCAAAATAGCAGCATTAATGCTTGAAGCATTAATGAGTGTTCAGTTGTCAC 4788
Db 4832 GACCTCTGCTTGTAGACTGGAACGACTGTCTCAGTCTGTTAAACAGGTTTGCAGCAGCAC 4891
Qy 4789 AGAGGATTAACCTGCTCAGGTCTGTATACAGAAACATCTGGTGTGCTTCTTATCAGTTG 4848
Db 4892 ATTGGAAGTGGCGTACTTCTCTGGTGTACAAAGGACACTTCTGTAAACTCTTTGCTATCCT 4951
Qy 4849 TTCTTCTTAACATAAATCAGAAATACACATGCTTCTGTCTATAGATCTGTTTCATG 4908
Db 4952 TGGTGCCTTAGTGAAGCTATGAAGACCTCCCGTGCATACAGACATCGACATGTTTCACT 5011
Qy 4909 TTTTGGTGGGTGCTGTGTTAGCAATTCCTTGTATTTGGGATGACCTCTGTGATCTGC 4968
Db 5012 TGCTGTGGGCTGTGCTGCTTTCAGCTCTGCTG-----TCAGG 5056
Qy 4969 AGCCTTCTCAGTTAGTTCTTCTTATACCACTTTTATCTCTTCAATTTGATACCAATGG 5028
Db 5057 ATTTTTCAGGAAGCAGCTGGCCACTGGGACCTGCGACATCTTCCACTTGTGTACCATGG 5116
Qy 5029 CACATGCTTACATCTACTTACAGTAGACACAGGCTTA---CCCCCTGCTCAGGTTTC 5085
Db 5117 CACATCTGACAGATCTTACTTACCTCAATGACAGAGAGAAATGGCATGGATCAAGAGA 5176
Qy 5086 AAGAAGCAGTGAAGAGGCTCATTTCCGCTCTTCTTTTGGCAGAAATTTCTCAATATA 5145
Db 5177 ATCCCACTGGGGAAGAGAACTGGCCATTTCTCTTTTGGCACAACACTTTCACCAATATA 5236

```
QY 5146 CAAAGTGGCTCCATTGGGTGTGATATTCCTGGCTGGTATTTGTGGTCTCACTGAAGAATG 5205
Db 5237 CTGGAAGTGCCTTGAAGAGCCCTCGGCTGGCACTGTGGAGGAGCGTCCGGGCGG 5296
QY 5206 GCATCACCCCTTATCTTCGCTGTGTGCAATGTTTTCACATATTTACTTGGGGTAACTC 5265
Db 5297 CCATCATGCTTTCTCCTCAAGTGTCTGTCTGTTTTCACATATTTAAATGGAGTCCGG 5356
QY 5266 GCGCTGAGAACTGATACCAATTCCTGAGAGGAGATACNGTGCATCTGTAGCTATC 5325
Db 5357 CCCCTCCAGACCTTCA---AGTTTCTGGAACAAGCCATTTTGAACACTTATGTAACTACC 5413
QY 5326 TATCTTTTACCTACAAATTTGTTCTGCTCTTCACAGAAATATTTGGGATACTGTAAAGCCCT 5385
Db 5414 TTTCTTACCAACCAACTCACTTACCTTTTTCAGAAACACNGTGCATCATGAATCCC 5473
QY 5386 TGCTCCAGAGTGTGTGCAGATCTGCTTACTTAAACTGTTTGAAGCAAAAAACACCG 5445
Db 5474 TGATTTGAAGTTGGTGCCAGAACAGTGAAGTTAAACGGTATCTAAATGGCGAGAGGAG 5533
QY 5446 TGGTCAGGTACCTAGAAAAGAAATAGTTTGTATAGAGCTTCCTGATGACTATAGCTGCC 5505
Db 5534 CGATAAGCTTACCCAGAGAGGACTAACAACTGTATAGACCTTTCCAGAGGATTTACAGCGCC 5593
QY 5506 TCCTCAATCAAGCTTCTCATTTTCAGTGTGCCACGGTCTGCAGATGATGAGCGAAGCATC 5565
Db 5594 TCATTTAACCAAGCATCCAAATTTCTGTCGCCCAATTCAGTGGCGACAGAGCAGAGCTC 5653
QY 5566 CTGCTCCTCTGCTTTTCTGTGGGGCTATATATGTTTCTCAGAAACATTTGCTGCCAGBAAA 5625
Db 5654 CTACTCTGTGCTCTGTGTGGAGTCTCCTCTGCTCTCAGAGTTTACTGCTGCCAAGCTG 5713
QY 5626 TTGTGAACGGGAAGAGGTGGAGCTTGCAATTTTTCAGCACTTCACCTGTGGAGCGGAG 5685
Db 5714 AGCTGGAGGTGAGGACGTCGAGCGCTGCACAGCACACACCTACTCTCTCGCGCTCCGGGG 5773
QY 5686 TCTGCATTTTCTTAAATTCAGAGATGCCAGTGTCTGCTGCTCAAGTAAAGCCAGAG 5745
Db 5774 CCGGCATCTTCTGAGAGTCCGGGAATGT/CAGTGCTATTTTGTACTGGCAAAACCAAG 5833
QY 5746 GCTGTGCTTATCCAGCTCCTTACTTGGATGAATATGGAGAAACAGACCCCTGCTGAAGA 5805
Db 5834 GATGTTTTTATCTCTCTCTACCTTGAGACTATGGAGAGACCGACAGGAGCTCAGAC 5893
QY 5806 GGGGCAACCCCTTCATTTATCTCTGAGCGGTATCGAAGCTCATTGTTGCTGGCAAC 5865
Db 5894 GAGGAAATCTTTTACATTTATGCCAAGAGCGGTTTTCGAAGATCCAGAAGCTCTGGCAGC 5953
QY 5866 RACACTGCATTTATAGAAGATTTGCTAGGAGCCAGAGACTAATCAGATGTTTATTTGGAT 5925
Db 5954 AGCATAGTATACAGAGGAGATCGGACAGCGCGCAGAGGCTTAACAGACCCCTGGTCGAA 6013
QY 5926 TCAACTGGCAGTTACTGTGAGCTC 5949
Db 6014 TTGACTGGCAGCATTTATAATCGC 6037
```

RESULT 13

```
US-10-755-889-634
; Sequence 634, Application US/10755889
; Publication No. US20040171823A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-kB
; FILE REFERENCE: D0284 NP
; CURRENT APPLICATION NUMBER: US/10/755,889
; PRIOR FILING DATE: 2004-01-13
; PRIOR FILING DATE: 2003-01-14
; PRIOR FILING DATE: 2003-05-12
; NUMBER OF SEQ ID NOS: 823
```

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 634

; LENGTH: 6158

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-755-889-634

Query Match 15.6%; Score 981; DB 19; Length 6158;

Best Local Similarity 55.9%; Pred. No. 3.8e-259; Indels 90; Gaps 12;

Matches 2157; Conservative 0; Mismatches 1610;

QY 2130 CTAAAGTATATCTCGATCAGCAAAACCCACATATATGGAACAGAAATTAAGAATGCAGTTC 2189

Db 1 CTCAAGTATGTGTTAAATTAGCAAAACCAACTGAATGGTCAGATGAGTCGAGCAGAAAGTTC 60

QY 2190 CTTGAAGGTTTTCGATCTTTTGAAGATCTTTTACCTGTATGACAGGGAATGGAAGAATC 2249

Db 61 CTAGAAGGGTTTGTATGCTTTTGGAAATTAATAAATGTATGACAGGGAATGGATCCAAAT 120

QY 2250 CGAAGACAGGTTGGGCAACACATTTGAAGTGCATCTGATTGGGAGGCTGCCATTTGCTATA 2309

Db 121 ACACGTCAAGTAGACCAACATATTTGAATGGAACACAGATGGGAAGCAGCTTCACACTA 180

QY 2310 CAGATGCAATTTGAAGAATATTTTACTCATGTTTCAAGAGTGGTGTGCTTGTGATGAAGAA 2369

Db 181 CAATGAAATTTAACACATGTCATTTCAATGATGACGAGCTGGTGTGCTTCAGATGAAAAA 240

QY 2370 CTCTTACTTGGGCTTATAAAGAATGTCAAAAGCTGTGATGAGTGCAGTACAGTTTC 2429

Db 241 GTGTTAATCGAAGCTTCAAGAAATGTCTGCTGTACTGATGCGAGTGTGTCATGGTGGTTAT 300

QY 2430 ATATCTAGTAGCAAGACAGT---AGTACAAATCGTGTGCACATAGTTTGGAAACAAAGTCC 2486

Db 301 ACTGATGTGGAACAGCCAAATCACAATAGCAATTTGTGACATTCAGTGGAAACTATCAGA 360

QY 2487 TACAGAGTATCTGAGGATCTTTGTAAGCATACATCTGCCACTCTCTAGGACCCCTTGTGGT 2546

Db 361 TACTGTGTTTCCCAAGAAAAAGTTAGCATTCACCTCCAGTTTCTCGCTACTTACTTGCAGGT 420

QY 2547 CTTCAATGACGTTTAAAGCAGGCTGGTGTCTGTTTCAAGACTGCATGAATTTGTGTCTTTT 2606

Db 421 TTACATGTTATTTAAAGCAAAAGTGAAGTGCATATAAATTTCCAGAGCTCCTACCTCTA 480

QY 2607 GAGGACTTTTCAAGTAGAGGTACTAGTGAATATCTTTTACGTGTGTGGTGTGGTGGCC 2666

Db 481 AGTGAACCTTAGCCCAACCATGTTGATGAACACCTCTTAGATGTCTTGTCTGTGGGCC 540

QY 2667 CAGGTTGTTGCTGAGATGTGGCGAAGAAATGGACTGTCTTATTTAGCCAGGTGTTTAT 2726

Db 541 CAAGTACATGCGGAAATGTGGAGAAGAAATGGGTTCTCTCTAGTAAACACAGATTTATTAC 600

QY 2727 TACCAAGATGTTAAGTCAGAGAAAGAAATGATGATAAGATATCATGCTTCAGATTT 2786

Db 601 TACCATAATGTGAATTCAGACGTCGAGATGTTTGAACAGGATGTAGTAAATGCTTCAGACA 660

QY 2787 GGTGCATCTTTTAAATGGATCCCAATAAGTTCTTGTGTACTTGGTACTTCAGAGGTATGAACCT 2846

Db 661 GGTGTCTCCATGATGATCCAAATCATTCTCTGATGATCATGCTCAGCGGCTTTGAACCT 720

QY 2847 GCCGAGGCTTTTAAACAGACCATATCTACAAAA-----GACCAG 2885

Db 721 TATCAGATTTTTCAGTACTCTCCAGACTATGGAAGAAAGATTTAGTCTCTGAGATTTACCCATAAG 780

QY 2886 GATTGTATTAACAATATATAATACACTAATAGAAGAAATGCTTCAGGTCCTCATCTATATT 2945

Db 781 GATGTTGTTTCAGCAGAAACAAATCTCTAATAGAGAAATGCTATACCTCATTTAATATGCTT 840

QY 2946 GTGGTGAAGCTTATGTACTCTGGAGTGGAAATGTGACCAAGAGAGGTCACAATGAGA 3005

Db 841 GTTGAGAGAGATTTAGTCTCTGGAGTTGGACAGGTAATGCTACAGATGAATCAAGCGA 900

QY 3006 GAAATCATTTCACTTGTCTTTGCAATTTGAACCCCAATGCCACAGTGCATTTGCCAAAAATTTA 3065

Db 901 GAGATTATCCATCAGTTGAGTATCAAGCCTATGGCTCATAGTGAATTGGTAAAGTCTTTA 960
Qy 3066 CTTGAGAAATGAATAATGAATCTGGCTTAGAGAAATGTCTATAAACAAAGTGGCCCAATTT 3125
Db 961 CCTGAAGATGAGAACAAAGAGACTGGCATGGAGAGTGAATCGAAGCAGTTGCCCATTTTC 1020
Qy 3126 AAGAAACCAAGGTGATCAGGCCATCGAGTTTATGAACCTAAAGATGAATCACTGAAAGAC 3185
Db 1021 AAGAAACCTGGATTAAACGAGCAGGCATGTATGAACCTGAACCCAGAAATGTGCCAAGAG 1080
Qy 3186 TTCAATATGTATCTTTTATCATTTACTTCCAAACCCAGCATAGCAAGCTGAAACATATGCAG 3245
Db 1081 TTCAACTGTATTCTTATCATCTTTTCAAGGCGAGAACAGTCCAGGCGAAGAGCGCA 1140
Qy 3246 AAGAAAGAGAGAAACAAAGAAACAAAGATGAAGCAATGGCCGCCACCACTCTCTGAA 3305
Db 1141 CGGAAATTTGAAAGACAAATAGAGAGATACAGCACTCCCACTCCGGTGTGGCTCCA 1200
Qy 3306 TTCTGCCCTGCTTTCAGCAAGTGTATTAACCTTCTCACTGTGATATCATGTATGTACATT 3365
Db 1201 TTCTGCCCTGCTTTCAGCAAGCTGGTTAAACATTTTGCAGTCAGATGTCTATGTGTGCATC 1260
Qy 3366 CTCAGACCGTATTTGAGCGGCAATAGACACAGATTTCTAATCTGTGGACCGAAGGATG 3425
Db 1261 ATGGGAAACATTTCTGCATGGGCTGTGGAAACATAATGGATATGCTGTGTCAGATGCCATG 1320
Qy 3426 CTCCTAAATGGCTTTTATATCTGGCATTTGGTTTACTAGAGAGAGCAACAGCTTCAA 3485
Db 1321 CTGCAAGGGGTGTACATTTAATTTGGCATGGCACTACAGAGAGAGAAACAAACATTTAGAG 1380
Qy 3486 AAGCTCTCTGAAG---AAGAACTAACATTTGATTTTATCATTAAGCTTCAAGNTTGGGA 3542
Db 1381 AATGTCAGGAAAGACATGTAGTAAACATTTTACCTTCACTCAGAGATATCAAAACCTGGT 1440
Qy 3543 AGTTCAGCCATCAAT-----ATACAAATGCTTTTGGAAAACTCAAAGGAATCCC 3593
Db 1441 GAAGCGCCAAAAATTTCTCTAGCATATCTAGCTATGTGGAAAACTACAAATGCTCCC 1500
Qy 3594 CAGTTAGAGGCGCAGAGGACATGATAAGCTGGATATCTTCAGATGTTTGACACAGTGAAG 3653
Db 1501 TACCTAGAGTCCCAAGACACATGATCGGTGGATATTGAACACTTTTAAATGCTGTAAA 1560
Qy 3654 CGATTAGAGAAAAATCTTTTAAATTTGTAGCAACCAATCAGGATCGGAATCTATTAG 3713
Db 1561 AAGATGAGGGAGA-----GTTCACTTACAGTCCCGTGGCAGAGACAGAGAGAACCATTA 1614
Qy 3714 AATGATGAGATTACTCATGATAAGAAAGACAGAACGAAAGAAAGCTGAAGCTGT 3773
Db 1615 ATGGAGAGAGTTTCAAGGCAAAAGACATGCTCAGATGTCTGAAATGCAGCGCATTTTATGAT 1674
Qy 3774 AGGCTACATCGCAGAGATCATGGCTCAGATGTCTGCGCTTACAGAAAACTTCAATTGAA 3833
Db 1675 AGACTGCCAGAGAAAGACATGCTCAGATGTCTGAAATGCAGCGCATTTTATGAT 1734
Qy 3834 ACTCATAAATCATGTATGACAAATACATCAGAAATGCCCTGGGAAAGAGATTCATATAG 3893
Db 1735 GAAACAAAGAACTCTTTTCAAGCAGACATTTAGAACTGGATGCTCAACCTCTGCTGTCTT 1794
Qy 3894 GAGGAGAGAGACCCAGCAGTCACTACTCTAGAAATGCTTTGGTCTTAAACGG 3953
Db 1795 GATCA-----TAGCCCTGTGGCTTTCAGATATGACATTTACAGCACTGGGTCCTCACAA 1848
Qy 3954 GGTCCATCTGTACTTGAAGAGAGGTGTGAGCTGCATCTTTGGCAAGAGAACAGGAG 4013
Db 1849 ACTCAGGTTCTGAAACAAAGCAATTCGTATCATATGATATTGTCTCAAGAGGAGCGA 1908
Qy 4014 GTGAAAAATAGAAAAATAATGCCATGTATATCGGCTGTGTCCAGAAATCTACTGCCCTTA 4073
Db 1909 GTTAAAGTGAAGAGCAGGCAATGGTCTTGGCAGCAATTTGTTCAGAGATCAACTGTATTA 1968
Qy 4074 ACCAGCAGAGGGAACCCATAGAACTCTCAGAGAGAGCCCTAGACCCACTTTTCATG 4133
Db 1969 TCAAAAAACAGAAAGTAAATTTTATTCAGAGATCCAGAAAAATAT---GATCCCATTTATCATG 2025

Qy 4134 GATCCAGACTTTGGCATATATGAACTTATATACAGGAAGCTGTGTGATGTAATGACGCGAGTG 4193
Db 2026 CACCCTGATCTGTCTTGTGTGAAACACACACTAGTAGCTGTGGGCAATATATGATGCCCAT 2085
Qy 4194 TGCTGGCAGAAAGTATTTTGAAGCTGTACAGCTGAGCTCTCAGCAGCGCATTCATGTTGAC 4253
Db 2086 TGTGGCAAGGATTTTGTATTCCTTCAAGCTAAAGAAACGAGGAGGCAACAGAGATTA 2145
Qy 4254 CTTT---TGACTTGGAAAGTGGAGATATCTTTGGCCTCTTTTGCAAACTC 4301
Db 2146 CGCTTACATACGAGCTATGATGTAGAAAACGAGAAATCTCTTTGGCCCTTTTGTGAATGC 2205
Qy 4302 CTGTGCAATACGTGTATCCCATTTATCTTTTGGAACTCTCAAAAGATTAACAGTGAAGAT 4361
Db 2206 TTGAGTAATACTGTTTATTTCC-----TCTGCTGCTTCTCTCCAAAGAAATATTTTAAACAC 2259
Qy 4362 CGAGATGCTCTTGTCTCAACTTTTGTACCTGGCAGCGTGGATACAGACTGTTCTTGCCAGA 4421
Db 2260 AGGTTAAATTTTTCAGACCAACCAATCTGACTCAGTGGATTTAGAAACAAATATCTCAGCA 2319
Qy 4422 ATATCAGGTTATTAATATTAAGACATCTTAAAGGAGAAAAACCCAAATTTCTTATTTTAT 4481
Db 2320 ATAAAGCATTTACAGTTTCTTAGGAAAGAGAAAGTACTCTCTAATAATGCTCTACAAAG 2379
Qy 4482 CAAGGAATGGAGATTTACTTTGGAGTTCCATTCCTCCTGAGTTTTTGGCGTTGAGTCT 4541
Db 2380 AATTGAGAAAAATGTGGATGAATTAACAGCTCCCTGAAGGGTTTCAGGCTCTGATTTTGGCTCT 2439
Qy 4542 TCGATTAAATATCAAAATAGCATCAAGAAATGGTTATTCTTTTGGCCCAACAATTTAT 4601
Db 2440 AAGATCCCTTATTTGAGAGCATAAAGAAATGCTAAACGACATTTTGGAACTCTCCTACCTAC 2499
Qy 4602 AGAATGGAATGAAAGTCCACCTGATGAAAGGATCTCTCGAGTCCCATGCTGACCTGG 4661
Db 2500 AAGTGGGACTAAGGTTTCATCCCAATGAGAGGATCTCTGTGTTCCCATATATGTTGG 2559
Qy 4662 AGCACTCGCTTTTCACTATCCAGGCAATTTGAAATCTATTGGGAGATGAAGAAACCT 4721
Db 2560 GGTAGTCGGGTACACCATCCAAAGCATAGAAAGAAATTTTGGATGATGAAGATAAACCA 2619
Qy 4722 CTGTTTGGAGCACTTCAAAATAGGCGCATTAATGCTCTGAAAGCATTAATGCAAGTTGCA 4781
Db 2620 TTGTTTGGTCTTCTTACCTTTCAGACTGGATGATGCTGTCTTAGTCTATTGACGAGATTTG 2679
Qy 4782 GTTGACAGAGATTTACCTGCTCTCAGGTCCTGTATACAGAAACATCTGCTCTCTCTA 4841
Db 2680 GCAGCACTGGAAGTGGCATCAGTTTTCAGTGGTGGAGGACATTTTGTAAACTTTT 2739
Qy 4842 TCAGTTGTTTCTTCAACATTAATAATCAGAGATACACCATGCTTCTGTCTATAGATCTG 4901
Db 2740 GCATCACTGGTGCCTTAATGACAGCCATGAGGAACTTCCATGCTATATAGATATTGACATG 2799
Qy 4902 TTTCAATGTTTGGTGGGTGCTGTGTAGCATTTCCCATCTTGTATTTGGAGTACCCCTGTT 4961
Db 2800 TTTCAATTTTGGTGGGTGCTGTGTGCTTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 2846
Qy 4962 GATCTGAGCTCTTCTTCACTGTTAGTTTCTTCTTATTAACACCTTTATCTCTTCCATTTGATC 5021
Db 2847 ---TCAGGATTTTTCAGGATCAGGCTTGGCTGCTGAGAGACTTCAATTTTCACTCTGTT 2904
Qy 5022 ACCATGGCAACATGCTTTTCAATATCTTACTTACAGTAGACACAGGCC---TACCCTTGTCT 5078
Db 2905 ACTATGGCACACATCATACAGATCTTACTTACCTCATGTACAGAGAGAAATGGCATGGAT 2964
Qy 5079 CAGTTTCAAGAGACAGTGAAGGCTCATTTCCGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5138
Db 2965 CAAGAAAATCCCTCTTGTGAAGAGAAATCAGCAGTCTTCTGCTTTGTATTAATAACACTTCA 3024
Qy 5139 CAATATCAAGTGGCTTCCATTTGGGTGTGATATTCTCTGCTGTGATTTTGTGGGTCTCACCTG 5198
Db 3025 CAGTATACGGGAAGTGCCTTGAAGAAATACCATCCGGCTGGCATCTCTGTGGAGGATGTC 3084

[illegible]

RESULT 14
US-09-529-063-56
; Sequence 56, Application US/09529063
; Patent No. US20020102542A1
; GENERAL INFORMATION:
; APPLICANT: FUKUSHIMA, DAIKICHI
; APPLICANT: SHIBAYAMA, SHIRO
; APPLICANT: TADA, HIDEAKI
; TITLE OF INVENTION: POLYPEPTIDE, CDNA ENCODING THE POLYPEPTIDE, AND USE OF
; TITLE OF INVENTION: THE BOTH
; FILE REFERENCE: Q58769
; CURRENT APPLICATION NUMBER: US/09/529,063
; CURRENT FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: PCT/JP98/04514
; PRIOR FILING DATE: 1998-10-06
; PRIOR APPLICATION NUMBER: JP 9-274674
; PRIOR FILING DATE: 1997-10-07
; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 56

```

; LENGTH: 3327
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-529-063-56

Query Match      12.9%   Score 813.6;   DB 9;   Length 3327;
Best Local Similarity 55.4%;   Pred. No. 5e-213;
Matches 1856;   Conservative 0;   Mismatches 1409;   Indels 87;   Gaps 11;

Qy  2632  TGGAAATACCTTTACGTTGTCTGGTGTGGTTGCCCGAGGTGTGTGCTGAGATGTGCGAA 2691
Db      8  TAGAACAACCCTCTTAGATGCTTGTGTTCTGTGTGCCCAAGTACATGCGGGAATGTGGAGAA 67

Qy  2692  GAAATGCACTGTCTCTTATTTAGCCAGGTGTTTTATTACCAAGATGTTTAAGTCGACAGAAG 2751
Db      68  GAAATGGGTTCTCTCTAGTAGTAACACGATTTATTACTACCAATAATGTGAATGCAGACGCTG 127

Qy  2752  AAATGTATGATAAAGATATCATCATGCTTCAGATTTGGTGCATCTTTTAATGGATCCCAATA 2811
Db  128  AGATGTTTGACAGGATGTAGTAAATGCTTCAGACAGGTGTCTCATGATGGATCCAAATC 187

Qy  2812  AGTTCCTTTACTGTGTACTTCAGAGGTATGAACCTGCCGAGGCTTTTAAACAAGACCAATAT 2871
Db  188  ATTTCCCTGATGATCATGCTCAGCGCGCTTTGAACTTTATCAGATTTTCAGTACTCCAGACT 247

Qy  2872  CTACAAA-----GACCAGAGTTTGATTAACAATATAATACAC 2910
Db  248  ATGGAAAAAGATTTAGTTCGTGAGATTAACCCATAGGATGTTGTTTCAGCAGAAACAATACTC 307

Qy  2911  TAATAGAGAGAAATGCTTCAGGTCTCATCTATATTGTGGGTGAGCGTATGTATACCTGGAG 2970
Db  308  TAATAGAGAGAAATGCTATACCTCATTAATATGCTTGTGTGAGAGAGATTTAGTCTCGGAG 367

Qy  2971  TGGAAATGTGAACAAAGAGAGGTCAATAGAGAGAAATCATTCACCTGCTTTGCATTTG 3030
Db  368  TTGCACAGGTAAATGCTTACAGATGMAATCAAGCGAGAGATTTATCCATCAGTCAGTATCA 427

Qy  3031  AACCCATGCCACACAGTGCCATTGCCAAAAATTTTACTCGAGATGAATAATTAATGAACCTG 3090
Db  428  AGCCATATGGCTCATAGTGAATTTGGTAAGTCTTTTACCTGAAGATGAGAACAGGAGACTG 487

Qy  3091  GCTTAGAGAAATGTCATAAAACAAATGGCCACATTTAAGAAACCAAGGTGTATCAGGCCATG 3150
Db  488  GCATGGAGAGTGTAAATCGAAGCACTTGCCCATTTCAAGAAACCTGGATTAACGAGACGAG 547

Qy  3151  GAGTTTATGAACATAAAGATGAATCACTGAAAGACTTCAATATGTACTTTTATCATTTACT 3210
Db  548  GCATGTATGAATGAACACCAAGATGTGCCAAAGAGTTCAACTTGATTTCTTATCAGCTTTT 607

Qy  3211  CCAAAACCCAGCATAGCAAGCGCTGACATATGCGAGAAAGAGGAGAAACACAGAAACA 3270
Db  608  CAAGGGCAGAACACTCCCAAGGCAGAGAACGCGCAACGGAATTTGAAAGAACAAATAGAG 667

Qy  3271  AAGATGAAGCAATTGC CGCCACACCACTCTCTGAATTTCTGCCCTGCTTTTCAGCAAACTGA 3330
Db  668  AAGATACAGCACTCCCACTCCGGTGTGCTCCATTTCTGCCCTCTGTTTGCAGACCTGG 727

Qy  3331  TTAACTTCTCAACTGTGATATCATGATGTACATTTCTCAGACCGTATTTGAGCGCGGCAA 3390
Db  728  TTAACTTTTGCAGTCAGATGTCAATGTTGTGTCATCATGGGAAACAAATCTGCAATTGGGCTG 787

Qy  3391  TAGACACAGATTTAACTTTGTGGACCGAAGGATGCTCCAAATGGCTTTTTCATATTTCTGG 3450
Db  788  TGGAAACATAATGGAATATGCTTGGTCAGATTCATGCTGCAAGGGTGTATACATTTAATTG 847

Qy  3451  CATTTGGGTTTACTAGAAAGAGAACCAACAGCTTCAAAAAGCTTCCTGAAAG---AAGAAGTAA 3507
Db  848  GCATGGCACTACAAGAGAGAAACAAACATTTAGAGATGTACAGGAAGAGCATGTACTGTAA 907

Qy  3508  CATTTGACTTTTATCATTAAGCTTTCAAGATTTGGGAGTTTCAGCCATGAAT-----A 3558
Db  908  CATTTACCTTTCACTCAGAGATATCAAAACCTGTGTGAAGCGCCAAAAAATTTCTCTAGCA 967

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 8, 2005, 18:43:15 ; Search time 17538 Seconds
(without alignments)
13690.809 Million cell updates/sec

Title: US-10-758-672A-1
Perfect score: 6308
Sequence: 1 gccagaattccgcagcagg.....aatttgatttggtgtttt 6308

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_hic: *
4: gb_est3: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_gsa1: *
9: gb_gsa2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2441.6	38.7	4386	9 AY405198	AY405198 Homo sapi
2	2334.2	37.0	4386	9 AY405199	AY405199 Pan trogl
3	2096	33.2	2561	3 AK089616	AK089616 Mus muscu
4	1671.4	26.5	4247	9 AY405200	AY405200 Mus muscu
5	826.6	13.1	360	4 BM472160	BM472160 AGENCOURT
6	806.4	12.8	3641	3 BC044903	BC044903 Mus muscu
7	771.4	12.2	865	5 BQ233617	BQ233617 AGENCOURT
8	740.2	11.7	3058	3 AK078173	AK078173 Mus muscu
9	723	11.5	838	1 AJ820333	AJ820333 J7820333
10	710.4	11.3	712	5 BX104087	BX104087 BX104087
11	687.8	10.9	3627	3 AK083320	AK083320 Mus muscu
12	674.6	10.7	682	1 AI929033	AI929033 au64c10.y
13	668.8	10.6	898	5 BU155953	BU155953 AGENCOURT
14	664.8	10.5	668	5 BU618516	BU618516 UI-H-FH1-
15	656.6	10.4	756	4 BG534574	BG534574 602553425
16	637.4	10.1	641	1 AI361043	AI361043 QY03f11.x
17	635.2	10.1	774	5 BU955331	BU955331 AGENCOURT
18	627.2	9.9	797	4 BG862813	BG862813 602799074
19	600	9.5	644	7 CN276943	CN276943 170005321
20	594.8	9.4	640	6 CD368666	CD368666 UI-H-FH1-
21	585.2	9.3	695	6 BY756516	BY756516 BY756516
22	583.6	9.3	703	4 BM114217	BM114217 L0801D04-
23	576	9.1	729	5 BU703363	BU703363 UI-M-FD0-
24	573.2	9.1	782	4 BI086469	BI086469 602849734

25	567.4	9.0	689	7 CN528381	CN528381 UI-M-HQ0-
26	560	8.9	560	6 CB159971	CB159971 K-ES02019
27	557.4	8.8	681	7 CK711118	CK711118 959357 WA
28	557	8.8	565	2 BF063405	BF063405 7h89406.x
29	553.4	8.8	567	7 CN276944	CN276944 170004241
30	553.2	8.8	730	7 CV116540	CV116540 AGENCOURT
31	551.4	8.7	778	7 CF741991	CF741991 UI-M-HB0-
32	545.4	8.6	684	6 BY764809	BY764809 BY764809
33	539	8.5	751	7 CF726597	CF726597 UI-M-HB0-
34	536	8.5	774	7 CK635421	CK635421 UI-M-HN0-
35	533.4	8.5	744	7 CP539210	CP539210 UI-M-GH0-
36	529.2	8.4	691	6 BY736138	BY736138 BY736138
37	521.2	8.3	528	5 BX119443	BX119443 BX119443
38	517.6	8.2	659	6 BY750548	BY750548 BY750548
39	512.2	8.1	606	2 BE589438	BE589438 195602 BA
40	511.4	8.1	752	6 CA313244	CA313244 UI-CF-FNO
41	509.4	8.1	633	2 BB622408	BB622408 BB622408
42	501.4	7.9	506	5 BX486919	BX486919 DKF2p686B
43	499.6	7.9	777	6 CB232501	CB232501 AGENCOURT
44	497.4	7.9	719	5 BU623096	BU623096 UI-H-FL1-
45	496	7.9	496	5 BX281065	BX281065 BX281065

ALIGNMENTS

RESULT 1
AY405198
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
SOURCE
gene
ORIGIN
Query Match
Best Local Similarity
Matches 3440; Conservative
QY
DB

AY405198 4386 bp DNA linear GSS 12-DEC-2003
Homo sapiens HCM2144 gene, VIRTUAL TRANSCRIPT, partial sequence,
Genomic survey sequence.
AY405198 GI:39761172
GSS.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 4386)
Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, P., Murphy, B.,
Perriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 4386)
Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, P., Murphy, B.,
Perriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
Location/Qualifiers
1..4386
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>4386
/locus_tag="HCM2144"

QY 756 ACCCTCAGCGTCTGGCATCTTGGTGGGATCAGCAAGTTGATTTTATCTACTGCTTTCTTG 815
DB 61 ACCCTCAGCGTCTGGCATCTTGGTGGGATCAGCAAGTTGATTTTATCTACTGCTTTCTTG 120
QY 816 CATCATTTGGCACAATTTGGTGGCAGAAAATTACTTTGCTGAAATGGACCCAGACTTGGAA 875
DB 121 CATCATTTGGCACAATTTGGTGGCAGAAAATTACTTTGCTGAAATGGACCCAGACTTGGAA 180
QY 876 AAGCAGGAGGAAGTGTAACAATGTCAATTAATTCATCTCCACTGGAATGGTACTTATTGGA 935
DB 181 AAGCAGGAGGAAGTGTAACAATGTCAATTAATTCATCTCCACTGGAATGGTACTTATTGGA 240
QY 936 GAAGATCCAGATATTTGCTTAGGAATTTGAAGACAGTGGAGCATTTTCAGCTTTTGTGG 995
DB 241 GAAGATCCAGATATTTGCTTAGGAATTTGAAGACAGTGGAGCATTTTCAGCTTTTGTGG 300
QY 996 AGGGTTTTCAAAAGTGGAGAGACAACCTTATTTCTTGAGGGATTTGTCAAAATTGATCCAA 1055
DB 301 AGGGTTTTCAAAAGTGGAGAGACAACCTTATTTCTTGAGGGATTTGTCAAAATTGATCCAA 360
QY 1056 TGTGTACTCTGTATGACTGCTTTCCAGGACAGTGTTCATATAAATCATCGTTTCAAGATG 1115
DB 361 TGTGTACTCTGTATGACTGCTTTCCAGGACAGTGTTCATATAAATCATCGTTTCAAGATG 420
QY 1116 CATACTTCTACTGGAGGGTCTGTGACTGTGGAGACACAGAGGCATGGAAAACCTGGC 1175
DB 421 CATACTTCTACTGGAGGGTCTGTGACTGTGGAGACACAGAGGCATGGAAAACCTGGC 480
QY 1176 CTTTTTGTGTAATCATGAACCTGGAAGACAGGTACTATATAAGAGAATTCACGCTGT 1235
DB 481 CTTTTTGTGTAATCATGAACCTGGAAGACAGGTACTATATAAGAGAATTCACGCTGT 540
QY 1236 CCGTTCGAATGAAGAGTAATTTGTCCAGGACAGAAAATATTTCTTCAGTGATAAATAT 1295
DB 541 CCGTTCGAATGAAGAGTAATTTGTCCAGGACAGAAAATATTTCTTCAGTGATAAATAT 600
QY 1296 GTCGTAGAATGATATATGGGAAGAGGAAAGAACTGCTCCTGAATCCAGATAAG 1355
DB 601 GTCGTAGAATGATATATGGGAAGAGGAAAGAACTGCTCCTGAATCCAGATAAG 660
QY 1356 GAGAAAATGAAGATACTATTGTGCTTTTCAATGATGAACACCATTCATATGACCA 1415
DB 661 GAGAAAATGAAGATACTATTGTGCTTTTCAATGATGAACACCATTCATATGACCA 720
QY 1416 GTCATATACAGCCTACAAAGAGCTCTTGACTGTGAGCTGCGAGGCCAGTTGCAATCC 1475
DB 721 GTCATATACAGCCTACAAAGAGCTCTTGACTGTGAGCTGCGAGGCCAGTTGCAATCC 780
QY 1476 ACTGCCATTGACAAAGAGGCTGTCGGGCTGTTAAAGCGGAGCTTATGCTGCTGCCAG 1535
DB 781 ACTGCCATTGACAAAGAGGCTGTCGGGCTGTTAAAGCGGAGCTTATGCTGCTGCCAG 840
QY 1536 GAAGCAAGGAAGATATAAGAGTCAATTCAGAAAATGTCTCAACATCCACTTCATGTA 1595
DB 841 GAAGCAAGGAAGATATAAGAGTCAATTCAGAAAATGTCTCAACATCCACTTCATGTA 900
QY 1596 GAAGTATTACATTCAGAGATATTGGCTCATFCAGAAAATTTGCTTTGGCTGTGGTCTCTGG 1655
DB 901 GAAGTATTACATTCAGAGATATTGGCTCATFCAGAAAATTTGCTTTGGCTGTGGTCTCTGG 960
QY 1656 ATGAACAAAATTTATGACTTATTCAGTGACTTTAGGCAGATCTTTTGGCCAGCATGCC 1715
DB 961 ATGAACAAAATTTATGACTTATTCAGTGACTTTAGGCAGATCTTTTGGCCAGCATGCC 1020
QY 1716 AGAAGAACTCAGCTCGAGAAATCCCTGTCTCATTAAGCAGGTTAAATGCTTTTCGGATGCA 1775
DB 1021 AGAAGAACTCAGCTCGAGAAATCCCTGTCTCATTAAGCAGGTTAAATGCTTTTCGGATGCA 1080
QY 1776 AAGCTTTTATAAGGTCGCCCTAAGATCTCTTCATGAATTTGATCTTTCAGCAGTTTTTTTATG 1835
DB 1081 AAGCTTTTATAAGGTTCCGCTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 1140

QY 1836 GAGATGGAATACAAAACTCTTTGCTATGGAATTTGTGAAGTATTATAAACACTGCAG 1895
DB 1141 NNN 1200
QY 1896 AAAGAAATATATCAGTGATGATCATGACAGAGTATCTCTATAAATGACACTTCAGTT 1955
DB 1201 NNN 1260
QY 1956 ATGTTTACTGTTCTCTACTCTGCTCGACATCTTATTGAAGAGCAGAAATGTTTATCTCTGTC 2015
DB 1261 NNN 1320
QY 2016 ATTACTGAAACTCTGCTAGAAAGTTTACCTGAGTACTTTGGACAGAAACAATAAAATCAAC 2075
DB 1321 ATTACTGAAACTCTGCTAGAAAGTTTACCTGAGTACTTTGGACAGAAACAATAAAATCAAC 1380
QY 2076 TTCCAGGGTTATAGCCAGGACAAATTTGGGAAGAGTATATGCAAGTAAATATGACCTTAAG 2135
DB 1381 TTCCAGGGTTATAGCCAGGACAAATTTGGGAAGAGTATATGCAAGTAAATATGACCTTAAN 1440
QY 2136 TATATCTGATCAGCAAAACCCACAATATGGACAGAAAGATTAAAGAAATGCAGTTCTCTGAA 2195
DB 1441 NNN 1500
QY 2196 GGTTTTTCGATCTTTTTCGAAGATCTTACCTGTATGACAGGGAATGGAAGAAATCCGAAGA 2255
DB 1501 NNN 1560
QY 2256 CAGGTTGGGCAACACATTTGAAGTGATCTGTGATGGGAGGCTGCCATTTGCTATACAGATG 2315
DB 1561 CAGGTTGGGCAACACATTTGAAGTGATCTGTGATGGGAGGCTGCCATTTGCTATACAGATG 1620
QY 2316 CAATTTGAAGAAATATTTTACTCATGTTTCAAGAGTGTGTGTTGTGATGGAAGAACTCTTA 2375
DB 1621 CAATTTGAAGAAATATTTTACTCATGTTTCAAGAGTGTGTGTTGTGATGATGATGATGAT 1680
QY 2376 CTGTGGCTTATAAAGAAATGTCAAAAGCTGTGATGAGTGCAGTACCAAGTTTTCATATCT 2435
DB 1681 NNNNTGGCTTATAAAGAAATGTCAAAAGCTGTGATGAGTGCAGTACCAAGTTTTCATATCT 1740
QY 2436 AGTAGAAGACAGTATGATCAATTCGTGTGGAATAGTGTGGAAAACAAAGTCTTACAGATTA 2495
DB 1741 AGTAGAAGACAGTATGATCAATTCGTGTGGAATAGTGTGGAAAACAAAGTCTTACAGATTA 1800
QY 2496 TCTGAGGATCTGTGATGACATATCTGCCACTCTCTAGGACCTTCTGCTGCTTCTCATGTA 2555
DB 1801 TCTGAGGATCTGTGATGACATATCTGCCACTCTCTAGGACCTTCTGCTGCTGCTGCTGCTGCT 1860
QY 2556 CGTTTAAAGCAGGCTGCGTGTCTTTTCAAGACTGCATGAATTTGTGTCTTTTGGAGCTTT 2615
DB 1861 NNN 1920
QY 2616 CAAGTAGAGGTAAGTGGAAATATCTTTACGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2675
DB 1921 CAAGTAGAGGTAAGTGGAAATATCTTTACGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980
QY 2676 GCTGAGATGCTGCGAAGAAATGAGTCTCTTATTAGCCAGTGTGTTTATTACCAAGAT 2735
DB 1981 GCTGAGATGCTGCGAAGAAATGAGTCTCTCTTATTAGCCAGTGTGTTTATTACCAAGAT 2040
QY 2736 GTTAAAGTGCAGAGAAAGAAATGATGATAAGATATCATCATGCTT----- 2780
DB 2041 GTTAAAGTGCAGAGAAAGAAATGATGATAAGATATCATCATGCTTCCAGAAAGAAATGATGAT 2100
QY 2781 -----CAGATTGCTGCACTTTTAAATGGATCCCAATAAGTTCTTGTGTACTG 2825
DB 2101 GCTGGATTTCTTTTCAGAGATTGGTGCATCTTTTAAATGGATCCCAANNNNNNNNNNNNNN 2160
QY 2826 GTACTTCAGAGGTATGAATCTTCCCGAGGCTTTTAAACAGACCATATCTACAAAGACCAG 2885
DB 2161 NNN 2220
QY 2886 GATTTGATTTAAACAAATATAATACACTAATAGAGAAATGCTTTCAGGCTCCTCATCTATAT 2945

Db 2221 |||||GATTCGATTAACCAATATATATACCTAAATAGAGAAATGCTTCAGGTCTCATCTATATT 2280
QY 2946 GTGGGTGAGCGTTATGTACCTGGAGTGGGAATGTGACCAAGAAGAGGTGCAAAATGAGA 3005
Db 2281 GTG----- 2283
QY 3006 GAAATCAITTCACCTTGCTTGGCATTGAAACCATGCCACACAGTGCCTATGCGCAAAATTTA 3065
Db 2284 ----- 2283
QY 3066 CCTGAGAATGAAATAATGAAACTGGCTTAGAGAAATGCTATAACAAGTGGCCACATTT 3125
Db 2284 ----- 2283
QY 3126 AAGAAACCAAGGTGTATCAGGCCATGAGTGTATGAACTAAAGATGAATCACTGAAAGAC 3185
Db 2284 ----- 2283
QY 3186 TTCAATATGTACTTTTATCATTTACTCTCAAAACCCAGCATAGCAGGCTGAACATATGCGAG 3245
Db 2284 ----- 2283
QY 3246 AAGAAAGGAGAAACAAGAAACAAGATGAGCAATTCGCCACCACCACTCTCGAA 3305
Db 2284 ----- 2310
QY 3306 TTCTGCCCTGCTTTTCAGCAAGTGATTAACCTTCTCAACTGTGTATCATGTATGATACATT 3365
Db 2311 TTCTGCCCTGCTTTTCAGCAAGTGATTAACCTTCTCAACTGTGTATCATGTATGATACATT 2370
QY 3366 CTCAGGACCGTATTTGAGCGGGCAATAGACACAGATTTCACTTTGGAACCGAAGGGATG 3425
Db 2371 CTCAGGACCGTATTTGAGCGGGCAATAGACACAGATTTCACTTTGGAACCGAAGGGATG 2430
QY 3426 CTCGAAATGCGCTTTTCATATCTGGCATTTGGGTTTACTAGAGAGAGCAACAGCTTCAA 3485
Db 2431 CTCGAAATGNN 2490
QY 3486 AAGAGCTCTCTGAAGAAGATTAACATTTGACTTTTATCATAGAGCTTCAAGATTTGGGAAGT 3545
Db 2491 NNN 2550
QY 3546 TCAGGCATGAATATACAAATGCTTTTGGAAAACTCAAGAAATCCCCAGTTAGAGGC 3605
Db 2551 NNN 2610
QY 3606 CAGAAGGACATGATAACGTGGATACCTTCAGATGTTTGAACACAGTGAAGCGATTAGAGAA 3665
Db 2611 NNN 2670
QY 3666 AAATCTTGTTTAATTGTAGCAACCATCAGGATCGGAATCTATTAAAGAAATGATGAGATT 3725
Db 2671 AAATCTTGTTTAATTGTAGCAACCATCAGGATCGGAATCTATTAAAGAAATGATGAGATT 2730
QY 3726 ACTCATGATAAAGAAAAAGCAGAACGAAAGAAAGCTGGAAGCTGTAGGCTACATCGC 3785
Db 2731 ACTCATGATAAAGAAAAAGCAGAACGAAAGAAAGCTGGAAGCTGTAGGCTACATCGC 2790
QY 3786 CAGAAGATCATGGCTCAGATGTCGTGCTTACAGAAAACTTCATTGAAACTCATAACTC 3845
Db 2791 CAGAAGATCATGGCTCAGATGTCGTGCTTACAGAAAACTTCATTGAAACTCATAACTC 2850
QY 3846 ATGTATGACAAATACATCAAAATGCTGCGGAAAGAGATTCCATTATGAGGAGAGAGC 3905
Db 2851 ATGTATGACAAATACATCAAAATGCTGCGGAAAGAGATTCCANNNNNNNNNNNNNNNC 2910
QY 3906 ACCCCAGCAGTCAGTGACTACTCTAGAAATGCTTTGGGTCTTAAACGGGGTCCATCTGTT 3965
Db 2911 ACCCCAGCAGTCAGTGACTACTCTAGAAATGCTTTGGGTCTTAAACGGGGTCCATCTGTT 2970
QY 3966 ACTGAAAGGAGGTGCTGACGTGCAATCTTTGCCAAGAGAAACAGAGAGGTGAATAATAGAA 4025
|||||

Db 2971 ACTGAAAGGAGGTGCTGACGTGCATCTTTGCCAAGAGAAACAGAGAGGTGAATAATAGAA 3030
QY 4026 AATAATGCCATGATATATCGGCTGTGTCCAGAAATCTACTGCTTTAAACCCAGCACAGG 4085
Db 3031 AATAATGCCATGATATATCGGCTGTGTCCAGAAATCTACTGCTTTAAACCCAGCACAGG 3090
QY 4086 GGAATAACCATAGAACTCTCAGGAGAGCCCTAGACCCACTTTTCATGGATCAGACTTGG 4145
Db 3091 GNN 3150
QY 4146 GCATATGGAACCTTATACAGGAAGCTGTGTCAATGTAATGACGACGCTGTGTGCGAGAAG 4205
Db 3151 GCATATGGAACCTTATACAGGAAGCTGTGTCAATGTAATGACGACGCTGTGTGCGAGAAG 3210
QY 4206 TATTTTGAAGCTGTACAGCTGAGCTCTCAGACGCGCAATCATGTGACCTTTTGTGACTTG 4265
Db 3211 TATTTTGAAGCTGTACAGCTGAGCTCTCAGACGCGCAATCATGTGACCTTTTGTGACTTG 3270
QY 4266 GAAAGTGGAGAAATATCTTTGCGCTCTTTGCAATCTCTGTGCAATCTGTGATGCCCAT 4325
Db 3271 GAAAGTGGAGAAATATCTTTGCGCTCTTTGCAATCTCTGTGCAATCTGTGATGCCCAT 3330
QY 4326 ATTCCTTTGCAACCTCAAAAGATATAACAGTGAGATGACAGATGCTCTTCTCAACTTTTG 4385
Db 3331 ATTCCTTTGCAACCTCA----- 3347
QY 4386 ACCCTGGCACGCTGGATACAGACTGTTCTGGCCAGAAATATCAGTTTATATATAAGACAT 4445
Db 3348 ----- 3347
QY 4446 GCTAAGAGGAAAAACCCAAATTCCTATTTCTTTAATCAAGGAATGGGAGATTCTACTTTG 4505
Db 3348 ----- 3347
QY 4506 GAGTTCCATTCCATCTCTGAGTTTGGCGTTGAGTCTTCGATTAAATATTAATAGCATC 4565
Db 3348 ----- 3381
QY 4566 AAGGAAATGGTTATTTCTCTTTGCCACAACTTTATAGAAATTTGGAATGGAAGTGCACCT 4625
Db 3382 AAGGAAATGGTTATTTCTCTTTGCCACAACTTTATAGAAATTTGGAATGGAAGTGCACCT 3441
QY 4626 GATGAAAGGATCTCTGAGTCTCCCATGTGACCTGGAGCACCTGCGCTTTTCAATATCCAG 4685
Db 3442 GATGAAAGGATCTCTGAGTCTCCCATGTGACCTGGAGCACCTGCGCTTTTCAATATCCAG 3501
QY 4686 GCAATTTGAAATCTATTTGGGAGATGAAAGAAACCTCTGTTTGGGAGCACTTCAAAATAGG 4745
Db 3502 GCAATTT-----GGTAAAGCTCATAGAAACCTGTTTGGAGGTT 3540
QY 4746 CAGCATAATGGTCTGAAAGCATTAATGAGTTTGCAGTTTGCAGAGGATTAACCTGTCTCT 4805
Db 3541 TCACATAATGGTCTGAAAGCATTAATGAGTTTGCAGTTTGCAGAGGATTAACCTGTCTCT 3600
QY 4806 CAGGCTCTGATACAGAAACATCTGGTCTTCTTCTATCAGTTGTTTCTTCTTAACATAAAA 4865
Db 3601 CAGGCTCTGATACAGAAACATCTGGTCTGCTCTCTAT----- 3637
QY 4866 TCAGAGATACACCACTGCTCTGTCTATAGATCTGTTTCTATGTTTGTGGGTGCTGTG 4925
Db 3638 ----- 3637
QY 4926 TTGATCTCCCATCTCTGTATTTGGATGACCTGTTGTATCTGCAGCCTTCTTCAGTTAGT 4985
Db 3638 ----- 3637
QY 4986 TCTTCTATAACCACTTTATCTCTTCCATTTGATCAACAAGGACACATGCTTCAGATA 5045
Db 3638 ----- 3637
QY 5046 CTACTTACAGTAGACACAGGCTACCCCTTGTGCTCAGGTTCAAGAAAGACAGTGAAGGCT 5105
Db 3638 -----CAGNN 3681

1356 GAGAAAAATGAAAGATACTATTGCTGCTCTTTCAATGATGAACACCACTTCATATGACCAC 1415
1357 |||||
661 GAGAAAAATGAAAGATACTATTGCTGCTCTTTCAATGATGAACACCACTTCATATGACCAC 720
662 |||||
1416 GTCATATACAGCCTACAAAGAGCTCTTGAGCTGTGAGCTGCGAGAGCCCAAGTTGCATACC 1475
1417 |||||
721 GTCATATACAGCCTACAAAGAGCTCTTGAGCTGTGAGCTGCGAGAGCCCAAGTTGCATACC 780
722 |||||
1476 ACTGCCATTGACAAAGAGGCTGTCGGGCTGTTAAAGCGGAGCTTATGCTGCTGCCAG 1535
1477 |||||
781 ACTGCCATAGACAAAGAGGCTGTCGAGCTGTTAAAGCGGAGCTTATGCTGTTGTCAG 840
782 |||||
1536 GAAGCAAAAGAGATATAAGAGCTCAATTCAGAAAATGCTCTCAACATCCACTTCATGTA 1595
1537 |||||
841 GAAGCAAAAGAGATATAAGAGCTCAATTCAGAAAATGCTCTCAACATCCACTTCATGTA 900
842 |||||
1596 GAAGTATTACACTCAGAGATTATGGCTCATCAGAAAATTTGCTTTGCGTCTTGCTCCTGG 1655
1597 |||||
901 GAAGTATTACACTCAGAGATTATGGCTCATCAGAAAATTTGCTTTACGCTCTGGTTCCTGG 960
902 |||||
1656 ATGAACAAAATTTATGAGCTATTCAAGTGACTTTAGCGAGATCTTTTGGCAAGCATGCCCTT 1715
1657 |||||
961 ATGAACAAAATTTATGAGCTATTCAAGTGACTTTAGCGAGATCTTTTGGCAAGCATGCCCTT 1020
962 |||||
1716 AGGAGAACCTGACTCGAGATCCCTGCTCTCATAGCAGGTAATGCTTTGGGATGCA 1775
1717 |||||
1021 AGAGAAAGAACCTGACTCGAGATCCCTGCTCTCATAGCAGGTTAATGCTTTGGGATGCA 1080
1022 |||||
1776 AAGCTTTATAAAGGTCGCCGTAAGATCTTCTCATGAATTCATCTTCAGCAGTCTTTTATG 1835
1777 |||||
1081 AAGCTTTATAAAGNN 1140
1082 |||||
1836 GAGATGGAATAAAAACTCTTTGCTATGGAATTTGTGAAGTATTTATAAACACTGCGAG 1895
1837 |||||
1141 NNN 1200
1142 |||||
1896 AAGAAATATATCAGTGATGATCATGACAGAGTATCTCTATACTGCATCTTCAGTTGAG 1955
1897 |||||
1201 NNN 1260
1202 |||||
1956 ATGTTTACTGTTCTACTCTGCTCGACATCTTATTGAAGAGCAGAGTGTATCTCTGTC 2015
1957 |||||
1261 NNN 1320
1262 |||||
2016 ATTACTGAAACTCTGCTAGAAGTTTACCTGAGTACTTGGACAGAAACAAATAAATTCAC 2075
2017 |||||
1321 ATTACTGAAACTCTGCTAGAAGTTTACCTGAGTACTTGGACAGAAACAAATAAATTCAC 1380
1322 |||||
2076 TTCCAGGTTATAGCAGGACAAATTTGGGAGAGTATATGACGTAATATGACCTAAAG 2135
2077 |||||
1381 TTCCAGGTTATAGCAGGACAAATTTGGGAGAGTATATGACGTAATATGACCTAAAN 1440
1382 |||||
2136 TATATCTCTGATCAGCAAAACCCACAATATGACAGAAAGATTAAAGATGCAATTCCTGAA 2195
2137 |||||
1441 NNN 1500
1442 |||||
2196 GGTGTTTCGATCTTTTGAAGATTTCTTACTGTATGAGGGAATGGAAGAAATCCGAAGA 2255
2197 |||||
1501 NNN 1560
1502 |||||
2256 CAGGTTGGGCAACACATTGAAGTGATCTCTGATGCGAGGCTGCCATGCTATACAGATG 2315
2257 |||||
1561 CAGGTTGGGCAACACATTGAAGTGATCTCTGATGCGAGGCTGCCATGCTATACAGATG 1620
1562 |||||
2316 CAATTGAAGAAATATTTTACTCATGTTTCCAGAGTGTGCTGTGATGATGAAGAACTCTTA 2375
2317 |||||
1621 CAATTGAAGAAATATTTTACTCATGTTTCCAGAGTGTGCTGTGATGATGATGATGATGAT 1680
1622 |||||
2376 CTGTGCTTATAAAGAAATGTCACAAAGCTGTGATGAGTGAGTGACGATGATGATGATGAT 2435
2377 |||||
1681 NNNNNNNNTTATAAGAAATGTCACAAAGCTGTGATGAGTGAGTGACGATGATGATGATGAT 1740
1682 |||||
2436 AGTAGCAAGACAGTAGTAGTAATCGTGTGGACATAGTCTTTGGAAACAAAGTCTTCACAGATG 2495
2437 |||||

1741 AGTAGCAAGACAGTAGTAGTCAATCGTGTGACATAGTTTGGAAACAAAGTCTTCACAGTA 1800
1742 |||||
2496 TCTGAGGATCTTTGTAAGCATATACCTGCGCACTCTCTAGGACCCCTGCTGCTTCATGTA 2555
2497 |||||
1801 TCTGAGGATCTTTGTAAGCATATACCTGCGCACTCTCTAGGACCCCTGCTGNNNNNNNN 1860
1802 |||||
2556 CGTTAAAGCAGGCTGGTCTGCTTTCAAGACTGCTGAATTTGCTGCTTTTGAAGACTTT 2615
2557 |||||
1861 NNN 1920
1862 |||||
2616 CAAGTAGAGGTACTAGTGGAAATATCCTTTACGTTGCTGCTGCTGCTGCTGCTGCTGCT 2675
2617 |||||
1921 NNNNNNNNGTACTAGTGGAAATATCCTTTACGTTGCTGCTGCTGCTGCTGCTGCTGCTG 1980
1922 |||||
2676 GCTGAGATGTGGCGAAGAAATGGAATGCTCTCTTTATTTAGCCAGGTGTTTATTACCAAGAT 2735
2677 |||||
1981 GCTGAGATGTGGCGAAGAAATGGAATGCTCTCTTTATTTANNNGTGTGTTTATTACCAAGAT 2040
1982 |||||
2736 GTTAAAGTCAGAGAAAGAAATGATGATAAAGATATCATCATGCTT----- 2780
2737 |||||
2041 GTTAAAGTCAGAGAAAGAAATGATGATAAAGATATCATCATGCTTCAGNNNNNTATGAT 2100
2042 |||||
2781 -----CAGATTGCTGCTCTTTAAATGGAATCCCAATAGTCTTGTGTTACTG 2825
2782 |||||
2101 GCTGGAATTTCTTTCACAGATTGGTGATCTCTTTTANNNGTGTGTTTATTACCAAGAT 2160
2102 |||||
2826 GTACTTCAGAGGTATGAACTTGCCGAGGCTTTTAAACAGACCATATCTCAAAAGACCAG 2885
2827 |||||
2161 NNN 2220
2162 |||||
2886 GATTGATTTAAACAATAATAACATACTAATAGAGAAATGCTTCAGTCTCTCATATATTT 2945
2887 |||||
2221 GATTGATTTAAACAATAATAACATACTAATAGAGAAATGCTTCAGTCTCTCATATATTT 2280
2222 |||||
2946 GTGGGTGAGCGTTATGTAACCTGGAGTGGGAAATGTGACCAAGAGAGGTCAATGAGA 3005
2947 |||||
2281 GTG----- 2283
2282 |||||
3006 GAAATCATTTCACTGCTGTTGCAATGAAACCATGCCACAGAGTGCATTCGCCAAAAATTTA 3065
3007 |||||
2284 ----- 2283
3066 COTGAGATGAAAAATAATGAAACTGGCTTTAGAGAAATGTCTAATAACAAAGTGGCCACATTT 3125
3067 |||||
2284 ----- 2283
3126 AAGAAACAGGTGTATCAGGCCATCGAGTTTATGAACCTTAAAGATGAATCACTGAAGAC 3185
3127 |||||
2284 ----- 2283
3186 TTCAATATGTACTTTTATCATTTCTCCAAACCCAGCATAGCAAGGTGCAACATATGCGAG 3245
3187 |||||
2284 ----- 2283
3246 AAGAAAGGAGAAAAACAAAGAAAAACAAAGATGAAGCAATTTGCCGCCACCACTCCTCGAA 3305
3247 |||||
2284 -----GCATTTGCCGCCACCACTCCTCGAA 2310
3306 TTCTGCCCTGCTTTTACGAGAAAGTGTATTAACCTTCTCAACTGTGATATCATGATGATGAT 3365
3307 |||||
2311 TTCTGCCCTGCTTTTACGAGAAAGTGTATTAACCTTCTCAACTGTGATATCATGATGATGAT 2370
2312 |||||
3366 CTCAGGACCGTATTTGAGCGGCAATAGACACAGATTTCTAACTTGTGACCCGAGGATG 3425
3367 |||||
2371 CTCAGGACCGTATTTGAGCGGCAATAGACACAGATTTCTAACTTGTGACCTGAGGATG 2430
2372 |||||
3426 CTCCAATGCGCTTTTTCATATTTCTGCAATGCGTCTTACTAGAAGAGAACACAGCTTCAA 3485
3427 |||||
2431 CTCCAATGNN 2490
2432 |||||
3486 AAGCTCTCTGAAGAGAAAGTAACTTTGACTTTTATCAATAGGCTTCAAGATTTGGGAAGT 3545
3487 |||||

QY 5706 AGAGATGCCGAGTGGTCTGGTTGAAGGTAAAGCCAGAGGCTGTGCTATCCAGCTCT 5765

Db 4147 NNN 4206

QY 5766 TACTTGGATGAATATGAGAAACAGACCCCTGGCTGAAGAGGCGCAACCCCTTCATT 5825

Db 4207 NNN 4266

QY 5826 TCTCGTAGCGGATCGGAAGCTCCATTGGTCTGGCAACAACACTGCATTATAGAAGAG 5885

Db 4267 TCTCGTAGCGGATCGGAAGCTCCATTGGTCTGGCAACAACACTGCATTATAGAAGAG 4326

QY 5886 ATTGCTAGAGGCAAGACTATCATGATTTATTTGGATTCACTGGCAGTTACTGTGA 5945

Db 4327 ATTGCTAGAGGCAAGACTATCATGATTTATTTGGATTCACTGGCAGTTACTGTGA 4386

RESULT 3

AK089616

LOCUS

DEFINITION

AK089616 2561 bp mRNA linear HTC 03-APR-2004

Mus musculus activated spleen cDNA, RIKEN full-length enriched library, clone:F830005C07 product:ubiquitin protein ligase E3 component n-recogin 1, full insert sequence.

ACCESSION

AK089616

VERSION

1 GI:26354612

KEYWORDS

HTC; CAP trapper.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 Carninci, P. and Hayashizaki, Y.

AUTHORS

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

TITLE

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL

Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE

20499374

PUBMED

11042159

REFERENCE

2

AUTHORS

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

TITLE

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL

Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE

20499374

PUBMED

11042159

REFERENCE

3

AUTHORS

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

TITLE

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer

JOURNAL

Genome Res. 10 (11), 1757-1771 (2000)

MEDLINE

20530913

PUBMED

11076861

REFERENCE

4

AUTHORS

The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.

TITLE

Functional annotation of a full-length mouse cDNA collection

JOURNAL

Nature 409, 685-690 (2001)

MEDLINE

11076861

PUBMED

11076861

REFERENCE

5

AUTHORS

The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL

Nature 420, 563-573 (2002)

MEDLINE

12000000

PUBMED

12000000

REFERENCE

6

AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, P., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,

Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.

Please visit our web site for further details.

URL: http://genome.gsc.riken.jp/

URL: http://fantom.gsc.riken.jp/

FEATURES

Location/Qualifiers

1..2561

/organism="Mus musculus"

/mol_type="mRNA"

/strain="NOD"

/db_xref="PANTOM DB:F830005C07"

/db_xref="taxon:10090"

/clone="F830005C07"

/tissue_type="activated spleen"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

14..>2560

/note="unnamed protein product; putative ubiquitin protein ligase E3 component n-recogin 1 (MGD) [MG:1277977, GB|NM_009461, evidence: BLASTN, 100%, match=2560]"

/codon_start=1

/protein_id="BAC40933.1"

/db_xref="GI:26354613"

/translation="NADKEMDGAERMDVSPPEPLAPORPASHMDQVDVFTAFHLHLA QLVPIYAEADMPDLKEESVQMSILPFLWYLFGEEDPDICLEKHKSGAQFOLCGKV FASGTTYSRCDAIDPTCVLDCDFQSSVHNHRYKMTSDPDICLEKHKSGAQFOLCGKV PFCVDHBEPRAGTTKESLHCPLENEVIAQARRIPFSVIYVEMTWESEKELPPELQ IREKNRYCVLFNDEHHSYDHYVLSQALDCLAEALHTTAIDKEGRRVAKGVY ATCQAKEDIKSHSNVSOHPLHVLHVMVMAHOKFALRLGSMWNKIMSYSSDPRQI FCOACLVEPGESENCLISRLMLWDKLYKGARKILHELIFSPFMEYKXKLPANEP VARYLQKQETISDDEHSISITALSQVMTPTVTLARHLIEQNVISVITELVLP EYLDNNKNFQSGQDKLGRVYVYICDLKYLISKVPIWTERLKAQFUEGFRSLKI LTCMQMEIRIQVQGHIEVDPDMAAIIQMLKNILMFQEWCAQDGLLVLVYKE CHKAVMRCSTNFMSTKSVOLCGHSLSTKSVSEDLVSIHLPLSRTLAGHLVLSR LGALSRLEFPVDFSDFOVEVLVYPLRCLVLAQVVAEMRRNGSLISQVFFYQDVK CREEMVYDILMLQIGASIMDPNKELLLVLRVLTDAFNKTSIKQDQLIKQNTLLI EMLQVLIYIVGVNTRREVMREIMHLCLEPMPHSAIARLNPENNET GLENVINKVATPKPKFVSGHGYELKDSLSLDFNMVYFHYSKTQHSKAEHM"

ORIGIN

Query Match 33.2%; Score 2096; DB 3; Length 2561;

Best Local Similarity 88.7%; Pred. No. 0;

Matches 2270; Conservative 0; Mismatches 290; Indels 0; Gaps 0;

QY 684 GTTTCCTCAAGATCGGACGAGGAGGCTGAGGAGTACTGAGAGATGGAATCAAGCGG 743

Db 2 GTTTCCTTAAGATCGGACGAGGAGGCTGAGGAGTACTGAGAGATGGAATCAAGCGG 61

QY 744 GAGTATCCCCAGACCCCTCAGCGTCTGGTGGATCAGCAAGTGAATTTTAT 803

Db 62 GAGCTCCCTGGCCCGCCAGCGCGCATCGTGTGGGATCAGCAAGTTGATTCTAT 121
Qy 804 ACTGCTTTCTGTCATCAATTTGGCACAATTTGGTGCCAGAAAATTTACTTTGCTGCAATGGAC 863
Db 122 ACTGCTTTCTTACATCAATTTGGCACAATTTAGTGCCAGAAAATTTATTTTGTGAGATGGAC 181
Qy 864 CCAGACTTTGGAAAGCAGGAGGAAAGTGTAACAAATGTCAAAATTCACATCCACTCGAAATGG 923
Db 182 CCAGATTTGGAAAGCAGGAGGAGTGTAACAGATGTCAATACTCACTCTTTGGAGTGG 241
Qy 924 TACTTATTTGGAGAGATCCAGATATTTGCTTTAGAGAAATTCAGACACAGTGGAGCATTT 983
Db 242 TACTTATTTGGAGAGATCCGATATTTGCTTTAGAGAAATTTAAACACAGTGGAGCGTTC 301
Qy 984 CAGCTTTTGGAGAGGTTTTCAAAAGTGGAGAGACAACCTATTTCTTTGAGGGATTTGTGCA 1043
Db 302 CAGTTGTGTGGAGGTTTTCAAAAGTGGAGAGACAACATATTTCTGTAGGGATTTGTGCA 361
Qy 1044 ATTGATCCAAATGTGTACTCTGTATGGAATCTTCCAGGACAGTGTTCATAAAATCAT 1103
Db 362 ATTGATCCAAATGTGTCTCTGTATGGAATCTTCCAAAGTAGTGTTCATAAAACCAT 421
Qy 1104 CGTTACAAGATCATACTTCTACTCGAGGAGGTTCTGTGACTGTGGAGACACAGAGGCA 1163
Db 422 CGTTACAAGATCATACTTCTACTCGAGGAGGTTCTGTGACTGTGGAGACACAGAGCG 481
Qy 1164 TGGAAAATCGGCCCTTTTGTGTAATCATGAACCTGGAGAGCAGGTTACTATAAAGAG 1223
Db 482 TGGAAAATCGGCCCTTTTGTGTGGATCACAGGCTTGGAGAGCAGGTTACTACAAAAGAG 541
Qy 1224 AATTCACGCTGTCCGTGTGAATCAAGAGGTAATTTGTCAGGACAGGAATATTTCTTCA 1283
Db 542 AGCTTACATTTGCCATTTGAATGAAGAGTGATTTGCTCAAGCCAGAGAAATATTTCCCTCG 601
Qy 1284 GTGATAAAATATGTCTAGAAATGACTATATGGGAAGGAAAAGAACTGCTCTCTGAA 1343
Db 602 GTGATAAAATATCATTTGTAGAAATGACTATATGGGAAGGAAAAGAAATTTGCTCTGAA 661
Qy 1344 CTCAGATAAGGAGAGAAAATGAAGATACTATTTGTCTTTTCAATGATGAACACCAT 1403
Db 662 CTCAGATAAGGAGAGAAAATGAACGATACTATTTGTCTTTTCAACGATGAGCACCAT 721
Qy 1404 TCATATGACCACTCATATACAGCTTACAAAGCTCTTGAAGTGTGAGCTCCAGAGGCG 1463
Db 722 TCGTATGATCATTTGATCTACAGTCTGACAGAGCTCTAGATTTGCGAGCTTGCAAGGCA 781
Qy 1464 CAGTTGCATACCACTGCCATTCGACAAAGAGGTCGTCCGGCTGTAAAGCGGAGCTTAT 1523
Db 782 CAGCTGCACACGCTGCCATCGACAAAGAGGTCGCGGGCTGTCAAAGCAGGTGTGTAT 841
Qy 1524 GCTGCTTGCAGGAAGCAAGAGATATAAAGAGTCAATTCAGAAATGTCTCTCAACAT 1583
Db 842 GCCACTTGCAGGAAGCAAGAGGATATAAAGAGTCACTCAGAGAACGCTCTCAGCAC 901
Qy 1584 CCACCTTCATGTAGAGTATTAACCTCAGAGATTTAGGCTCATCAGAAAATTTGCTTTGCGT 1643
Db 902 CCCCTTCATTTGGAGTGTGCACTCCGTTATAGGCTCACAGAAAATTCGCTCTGCGC 961
Qy 1644 CTTGGTTCTCGATGAACAAAATTTAGCTATTCAGGTGACTTTTAGGAGATCTTTTGC 1703
Db 962 CTTGGCTCTGATGAACAAAATTTAGCTATTCAGGTGACTTTTAGACAGATATTTTGC 1021
Qy 1704 CAAAGATGCTTAGAGAGAACTGTACTCGGAGATCCCTGTCTCATAAAGCAGGTTAATG 1763
Db 1022 CAGGCTGCTCGTAGAGAACTGTGCTCTGAAAATCCCTGTCTTATAAGCAGACTAATG 1081
Qy 1764 CTTTGGGATGCAAGCTTTATAAGGTGCCGTAGATCTTCATCAATTTGATCTTCAGC 1823
Db 1082 CTTTGGGATGCAAACTTTATAAGGTGCCGTAGATCTTCATCAATTTGATCTTCAGT 1141
Qy 1824 AGTTTTTTTATGGAGATGGAATACAAAACCTCTTTGCTATGGAATTTGTGAGATTTAT 1883
Db 1142 AGTTTTTTTATGGAGATGGAATACAAAACCTCTTTGCTATGGAATTTGTGAGATTTAT 1201

Qy 1884 AAACAACCTGCAGAAAGATATATCAGTGTATCATATGACAGAGTATCTCTATTAACATGCA 1943
Db 1202 AAACAACCTGCAGAAAGATATCATCAGCGACACACAGAGAGAGCAATCTCCATAACCGCC 1261
Qy 1944 CTTTCAGTTCCAGATGTTTACTTCTTCTACTCTGGCTCGACATCTTTATTTGAAGAGCAGAA 2003
Db 1262 CTGTCCGTGCAGATGCTCACCGTCCGACCTTGGCCCGGCATCTTTATTTGAAGAGCAGAA 1321
Qy 2004 GTTATCTCTGTCAATTAATGAAACTCTGTCTAGAGTTTACCTGAGTACTTTGGACAGGAA 2063
Db 1322 GTTATTTCTGTCAATTAATGAAACGCTGTCTAGAGTTTACCTGATATCTTTGGACAGGAA 1381
Qy 2064 AATAAATTTCAACTTCCAGGTTTATAGCCAGACAAAATTTGGGAAGAGTATATGACATTA 2123
Db 1382 AATAAATTTCAACTTCCAGGTTTATAGCCAGACAAAATTTGGGAAGAGTCTACGAGTTATA 1441
Qy 2124 TGTGACCTTAAAGTATATCTCTGATCAGCAAAACCCCAATATGACACAGAAAGATTAAGAATG 2183
Db 1442 TGTGACCTTAAAGTATATCTCTGATTAGCAAGCTGTCTATATGACACAGAAAGATTAAGAAGCG 1501
Qy 2184 CAGTTCTCTGAAAGTTTTCGATCTTTTGTGAAGATTTCTTACCTGTATGACAGGAAATGAA 2243
Db 1502 CAGTTCTCTGAAAGTTTTCGCTCTTTTCTGAAGATTTCTTACCTGTATGACAGGAAATGAA 1561
Qy 2244 GAAATCCCAAGACAGGTTTGGCAACACATTTGAAGTGAATCTTGAATTTGGAGGCTGCCAT 2303
Db 1562 GAAATCCCAAGACAGGTTTGGCAACACATTTGAAGTGAATCTTGAATTTGGAGGCTGCCAT 1621
Qy 2304 GCTATACAGATGCAATTTCAAGAAATTTTATCTCATTTTCAAGAGTGTGTCTTTGTGAT 2363
Db 1622 GCTATACAGATGCAATTTCAAGAAATTTTATTTGCTCATTTTCAAGAGTGTGTCTTTGTGAT 1681
Qy 2364 GAAGAACTTCTTACTTGTGGCTTTAAGAAATGTCAAAAGCTGTGAATTTGGAAACAAAG 2423
Db 1682 GAAGAACTTCTTACTTGTGGCTTTAAGAAATGTCAAAAGCTGTGAATTTGGAAACAAAG 1741
Qy 2424 AGTTTCATATCTGTAGACAGACAGTACTCAATCGTGTGGACATAGTTTGGAAACAAAG 2483
Db 1742 AATTTCAATCTGTAGACAGACAGTACTCAATCGTGTGGACATAGTTTGGAAACAAAG 1801
Qy 2484 TCCTACAGATATCTGAGGATCTTGTAAAGCATACATCTGCCACTCTCTAGGACCTTGT 2543
Db 1802 TCCTACAGATATCTGAGGATCTTGTAAAGCATACATCTGCCACTCTCTAGGACCTTGT 1861
Qy 2544 GGTCTTCATGTACGTTTAAAGCAGGCTGGTGTCTTTCAAGCTGATGAATTTGTGTCT 2603
Db 1862 GGTCTTCATGTACGTTTAAAGCAGACTAGGTGTCTATTTCAAGACTGATGAATTTGTGTCT 1921
Qy 2604 TTTTGAGGACTTTCAAGTAGAGTACTAGTGGAAATATCTTTTACGTTGTCTGTGTGT 2663
Db 1922 TTTTGAGGACTTTCAAGTAGAGTACTAGTGGAGTACCCGCTGCTGCTGTGTGTGT 1981
Qy 2664 GCCCAGGTTTGTGTGATGTGGCGAAGAAATGGAATGACTGTCTTTATTTAGCCAGGTTT 2723
Db 1982 GCTCAGGTTTGTGTGATGTGGCGAAGAAACGGCTCTCTCACTCATCAGCCAGGTTTTC 2041
Qy 2724 TATTAACCAAGATTTAAGTGCAGAGAGAAATGTATGATAAAGATATCATCTGTCTAG 2783
Db 2042 TATTAACCAAGATTTAAGTGCAGGAGGAAATGTACGATAAAGATATCATCTGTCTAG 2101
Qy 2784 ATTTGCTGCACTTTTAAATGATCCCAATTAAGTTCTTTTGTACTGTCTTCTCAGAGGTTAGAA 2843
Db 2102 ATTTGAGCATCTTAAATGATCCCAATTAAGTTCTTTTGTACTGTCTTCTCAGAGATAGAA 2161
Qy 2844 CTTGCCGAGGCTTTTAAACAAGACCATATCTACAAAAGACCAAGATTTTGTATTAACAATAT 2903
Db 2162 CTTACTGTGCTTTTAAACAAGACCATATCTACAAAAGACCAAGATTTTGTATTAACAATAT 2221
Qy 2904 AATAACATTAAGAGAAATGCTTCAGGTCCTCATCTATATTTGTGGTGAGGCTTATGTA 2963
Db 2222 AATAACATTAAGAGAAATGCTTCAGGTCCTCATCTATATTTGTGGGAGAACGTTATGTA 2281

```
QY 2964 CCTGGAGTGGGAATGTGACCAAGAGAGAGGTACAAATGAGAGAAATCATTCACTGCTT 3023
Db 2282 CCTGGAGTGGGAATGTGTACAGAGAGAGAGGTATATATGAGAGAGATTCTCACTTACTT 2341
QY 3024 TGCATTGAACCCCATGCCACACAGTGCATTTGCCAAAAATTTTACCTGAGAAATGAATAAT 3083
Db 2342 TGCATTGAGCCCATGCCACACAGTGCATTTGCCAGAAACCTTACTCTGAGACGAAATAT 2401
QY 3084 GAAACTGGCTTAGAGAAATGTCAATAACAAAGTGGCCACATTTAAGAAACAGGTGTATCA 3143
Db 2402 GAAACTGGCTTAGAGAAATGTCAATAACAAAGTGGCCACATTTAAGAAACAGGTGTATCA 2461
QY 3144 GSCCATGGAGTGTATGAATCACTGAAGATGAATCACTGAAGACTTCAATATGTACTTTAT 3203
Db 2462 GSCCATGGAGTGTATGAATCACTGAAGATGAATCACTGAAGACTTCAATATGTACTTTAT 2521
QY 3204 CATTACTTCCAAAACCCAGCATAGCAAGCTGAAACATATGC 3243
Db 2522 CATTACTTCAAAACACAGCATAGCAAGCTGAAACATATGC 2561

RESULT 4
LOCUS AY405200
DEFINITION Mus musculus HCM2144 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY405200
VERSION AY405200.1 GI:39761174
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 4247)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,E., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 4247)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,E., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source Location/Qualifiers
1..4247
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>4247
/locus_tag="HCM2144"

ORIGIN
Query Match 26.5%; Score 1671.4; DB 9; Length 4247;
Best Local Similarity 58.2%; Pred. No. 0;
Matches 3055; Conservative 0; Mismatches 1192; Indels 1003; Gaps 11;

QY 696 ATGCGCGAGAGAGAGCTGAGTACTGAGAGATGGAATCAGCGCGAGGTACCCAG 755
Db 1 ATGCGCGAGAGAGATGAGCGCGCGAGAGATGAGCGTGCAGCGCGAGCGCTCCCTG 60
QY 756 ACCCTTCAGCGTCTGCGATCTTGGTGGATCAGCAAGTTGATTTTATCTGCTTCTTG 815
Db 61 GCCCGCGAGCGCGCGATCGTGGTGGATCAGCAAGTTGATTTTATCTGCTTCTTCA 120
```

```
QY 816 CATCATTTGGCACAATTGGTGCCAGAAATTTACTTGTGTAATGACCCAGACTTGGAA 875
Db 121 CATCATTTGGCACAATTAGTGCCAGAAATTTATTTTGTCTGAGATGGACCCAGATTGGAA 180
QY 876 AACGAGGAGAAAGTGTACAAATGTCAATATTCACCTCCACTGGAATGGTACTTATTGGGA 935
Db 181 AACGAGGAGAGAGTGTACAGATGTCAATCTCACTCTTTGGAGTGGTACTTATTTTGGGA 240
QY 936 GAAGATCCAGATATTTGCTTTAGAGAAATTTGAAGCACAGTGGAGCATTTTCAGCTTTTGGG 995
Db 241 GAGGATCCGATATTTGCTTTAGAGAAATTTAAACACACAGTGGAGCGTTCCAGTTGTGGG 300
QY 996 AGGGTTTCAAAGTGGAGAGACAACTATTCTTGAGGGATTTGTGCAATTTGATCAACA 1055
Db 301 AAGGTTTTCAAAGTGGAGAAACAAATATTTCTGTAGGGATTTGTCAATTTGATCAACAG 360
QY 1056 TGTGTACTCTGTATGACTGCTTCCAGGACAGTGTTCATAAAATCATCTGTACAAGATG 1115
Db 361 TGTGTCTCTGTATGACTGCTTCCAAAGTAGTGTTCATAAAACCATCGTTTACAAGATG 420
QY 1116 CATACTTCTA CTGGAGGAGGTTCTGTGACTGTGGAGACACAGAGGCATGGAATCTGGC 1175
Db 421 CATACTTCTA CTGGAGGAGGTTCTGTGACTGTGGAGACACAGAGGCATGGAATCTGGC 480
QY 1176 CTTTTTGTGTAAATCATGAACTCTGGAAGACAGGTACTATATAAGAGAAATTCACGCTGT 1235
Db 481 CTTTTTGTGTGATCAAGAGCTCTGGAAGACAGGTACTATACAAAGAGAGCTTACATTTG 540
QY 1236 CGGTGAATGAGAGGTAAATGTCTCAAGCCAGGAAATATTTCTTCAGTGATAAATAT 1295
Db 541 CCATTAATGAGAGGTAAATGTCTCAAGCCAGGAAATATTTCTTCGTTGATAAATAT 600
QY 1296 GTCTGAGAATGACTATATGGAAGAGGAAAGAACTGCTCTGAACTCCAGATAAGG 1355
Db 601 ATTGTAGAAATGACTATATGGAAGAGGAAAGAAATGCTCTGAACTCCAGATAAGG 660
QY 1356 GAGAAATGAAAGATCTATTGTGTCTTTTCAATGATGAACACCATTCATATGACCAC 1415
Db 661 GAGAAATGAAAGATCTATTGTGTCTTTTCAAGATGAGCAACCATTCATATGATCAT 720
QY 1416 GTCATATACAGCTTACAAAGAGCTCTTGTACTGTGAGCTCGCAGAGCCAGTTGCATACC 1475
Db 721 GTGATCTAAGTCTGACAGAGCTCTAGATGTCGAGCTTGCAGAGACACAGCTGCACAG 780
QY 1476 ACTGCCATTGACAAAGAGGCTCGTGGGCTGTATAAGCGGAGCTTATGCTGTGCCAG 1535
Db 781 ACTGCCATCGACAAAGAGGCTCGCGGCTGTCAAGCAGGTGTGTATGCCACTTGCAG 840
QY 1536 GAAGCAAGGAGATATAAAGAGTCAATCAGAAATGTCTCTCAACATCCACTTCATGTA 1595
Db 841 GAAGCAAGGAGATATAAAGAGTCAATCAGAGAACGTCTCTCAGCAACCCCTCCATGTG 900
QY 1596 GAAGTATTACATCTCAGAGATTTATGGCTCATCAGAAATTTGCTTTGGCTTGTGGTCTCG 1655
Db 901 GAAGTGTGCTCACTCGGTGTTATGGCTCACAGAAATTTGCTCTGCGCTTGGCTCTCG 960
QY 1656 ATGAACAAAATTTATGAGCTATTCAAGTGAATTTTAGGCAGATCTTTTCCCAAGCATGCC 1715
Db 961 ATGAACAAAATTTATGAGCTATTCAAGTGAATTTTAGGCAGATTTTTCAGCGGCTGCC 1020
QY 1716 AGAGAGAAACCTGACTCGGAGAAATCCCTGTCTCATAAGCAGGTTAATGCTTTGGGATGA 1775
Db 1021 GTAGAAGAACCTGGCTCTGAAATCCCTGTCTTATAAGCAGACTAATGCTTTGGGATGCA 1080
QY 1776 AAGCTTTATAAGGTGCCCGTAAGATCCTTCATGAATTCATCTCAGCAGTGTTCCTTATG 1835
Db 1081 AAACTTTATAAGGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1140
QY 1836 GAGATGGAATACAAAACAACTCTTTGTCTATGGAATTTGTGAAGTATTTATAAACAACTGCAG 1895
Db 1141 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1200
QY 1896 AAAGAATATATCAGTGATGATCATGACAGAGATATCTCTATAACTGSCACTTTTCAGTTTCAG 1955
```


cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12328 row: C column: 08
High quality sequence stop: 720.
Location/Qualifiers
1. 960
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5576383"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 92"
/notes="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

FEATURES

source

ORIGIN

Query Match 13.1%; Score 826.6; DB 4; Length 960;
Best Local Similarity 96.1%; Pred. No. 5.9e-208;
Matches 890; Conservative 0; Mismatches 30; Indels 6; Gaps 4;

QY 4924 TGTAGCATTCCTGTTATGGAGTACCCCTGTTGATCTGCGACCTTCTTCAGTTA 4983
DB 1 TGTAGCATTCCTGTTATGGAGTACCCCTGTTGATCTGCGACCTTCTTCAGTTA 60

QY 4984 GTTCTTCTATACCACTTATCTCTTCATTTGATCACCAGGACATGCTTCAGA 5043
DB 61 GTTCTTCTATACCACTTATCTCTTCATTTGATCACCAGGACATGCTTCAGA 120

QY 5044 TACTACTTACAGTAGACACAGGCTACCCCTGCTCAGGTTCAAGAAGACAGTGAAGAGG 5103
DB 121 TACTACTTACAGTAGACACAGGCTACCCCTGCTCAGGTTCAAGAAGACAGTGAAGAGG 180

QY 5104 CTCATTCGCGATCTTCTTTTTCAGAAATTTCTCAATATACAGTGGCTCATTGGGT 5163
DB 181 CTCATTCGCGATCTTCTTTTTCAGAAATTTCTCAATATACAGTGGCTCATTGGGT 240

QY 5164 GTGATATCTCGCTGTTATTTCTGCTGCTCAGTGAAGATGCATCACCCTTATCTTC 5223
DB 241 GTGATATCTCGCTGTTATTTCTGCTGCTCAGTGAAGATGCATCACCCTTATCTTC 300

QY 5224 GCTGTGCTGCTATTTTTCACATATTTACTTGGGGTAACTCGGCTGAGGAACCTGCATA 5283
DB 301 GCTGTGCTGCTATTTTTCACATATTTACTTGGGGTAACTCGGCTGAGGAACCTGCATA 360

QY 5284 CCAATTCGAGAGGAGATACAGTGCACTCTGTAGTATCTATCTTTACCTACAAAT 5343
DB 361 CCAATTCGAGAGGAGATACAGTGCACTCTGTAGTATCTATCTTTACCTACAAAT 420

QY 5344 TGTCTCTGCTTCCAGGATATGGATCTGTTAGGCTTCTGAGGCTTCTCCAGGCTGGTG 5403
DB 421 TGTCTCTGCTTCTCCAGGATATTTGGATCTGTTAGGCTTCTGCTCCAGGCTGGTG 480

QY 5404 CAGATCTGCTTACTTAACTGTTGAAGCAAAAAACACCGTGGTCAAGTACCTTAGAA 5463
DB 481 CAGATCTGCTTACTTAACTGTTGAAGCAAAAAACACCGTGGTCAAGTACCTTAGAA 540

QY 5464 AAGAAATAGTTGATAGCTTCTGATGATATAGTCTGCTGCTGAATCAAGCTTCTC 5523
DB 541 AAGAAATAGTTGATAGCTTCTGATGATATAGTCTGCTGCTGAATCAAGCTTCTC 600

QY 5524 ATTTAGGTCCTCCAGCTCTGCAGATGATGAGGGAAGATCTCTGCTCTGCTTTCT 5583
DB 601 ATTTAGGTCCTCCAGCTCTGCAGATGATGAGGGAAGATCTCTGCTCTGCTTTCT 660

QY 5584 GTGGGCTATATGTTCTCAGAAATTTGCTGCCAGGAAATTTGTGAACGGGAGAGG 5643

DB 661 GTGGGCTATATGTTCTCAGAAATTTGCTGCCAGGAAATTTGTGAACGGGAGAGG 720

QY 5644 TTGGAGCTTGCACTTTTTCACGCACCTTCACTGTGGAGCCGAGCTCTGCAATTTTCTTAAAAA 5703
DB 721 TTGGAGCTTGCACTTTTTCACGCACCTTCACTGTGGAGCCGAGCTCTGCAATTTTCTTAAAAA 780

QY 5704 TCAGAGAATGCCAGAGTGG-TCTGTGTGAAGGTAAAGCCAGAGGCTGT-GCTTATCCAGC 5761
DB 781 TCAGAGAATGCCAGAGTGGTCTGTGTGAAGGTAAAGCCAGAGGCTGTGCTTATCCAGC 840

QY 5762 T-CTTACTTGGATGAATATGAGAA---ACAGACCTTGGCTGAAGAGGGGCAACCC 5817
DB 841 TCCCTTACTTGGGATGATATGAAAAAACAAGAACCTTGGGCTGAAGAGGGGGCAAAACC 900

QY 5818 TTCATTTATCTCGTGAGCGGTATCG 5843
DB 901 CCGTTTCATTTATCTTCTCGGAGCGG 926

RESULT 6
BC044903
LOCUS
DEFINITION
BC044903
ACCESSION
VERSION
BC044903.1 GI:28279967
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Mus musculus
REFERENCE
1 (bases 1 to 3641)
AUTHORS
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L.,
Schetz T.E., Brownstein M.J., Udén T.B., Toshiyuki S.,
Carninci P., Frange C., Raha S.S., Loquellano N.A., Peters G.J.,
McKernan R.D., Mullahy S.J., Bosak S.A., McEwan P.J.,
Abramson K.J., Malek J.A., Gunaratne P.H., Richards S.,
Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahey J., Helton E., Kettman M., Madan A., Rodrigues S.,
Sánchez A., Whiting M., Madan A., Young A.C., Shevchenko Y.,
Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D.,
Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S., Krzywinski M.I., Skalska U., Smalish D.E.,
Schnerch A., Schein J.E., Jones S.J. and Marra M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
PUBMED
12477932
REFERENCE
2 (bases 1 to 3641)
AUTHORS
Strausberg R.
Direct Submission
TITLE
Submitted (23-JAN-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK
COMMENT
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu

Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 40 Row: b Column: 10
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 22122576
 This clone has the following problem: frame shifted.

FEATURES

source

1. .3641
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="Czech II"
 /db_xref="taxon:10090"
 /clone="IMAGE:3493115"
 /tissue_type="Mammary tumor metastasized to lung. Tumor arose spontaneously from a senescent normal mammary (clonal) outgrowth infected with the virus MMTV."
 /clone_1ib="NCI_CGAP_bu29"
 /lab_host="DH10B"
 /notes="Vector: pCMV-SPORT6"

ORIGIN

Query Match 12.8%; Score 806.4; DB 3; Length 3641;
 Best Local Similarity 54.7%; Pred. No. 2.2e-202;
 Matches 1934; Conservative 0; Mismatches 1516; Indels 88; Gaps 13;

QY 2460 TGTGGACATAGTTTGGAAACAAAGTCCTACAGAGTATCTGAGGATCTTGAGCATACAT 2519
 DB 22 TGTGGACATCGGTGGAAACCATCAGATACGTGTGTTCCCAAGAAAAGTTAGCATTCAC 81

QY 2520 CTGCCACTCTAGGACCCCTTGCTGCTTTCATGTAGCTTTAAGCAGGCTGGGTGCTTT 2579
 DB 82 CTCCCAATTTCTGCTTGTCTTGAGGTTTGCAATGTTTAAAGCAAAAGTGAAGTGCCA 141

QY 2580 TCAGACTGATGAATTTGTGCTTTTGAGGACTTTCAAGTAGAGTACTAGTGGAAATAT 2639
 DB 142 TATAAATTTCCAGAGCTCCTACCTTAAGTGAACCTGAGCCACCCATGTTGTAGAGCAT 201

QY 2640 CCTTTAGTGTCTGCTGTTGGTGGCCAGGTTGTTGCTGAGATGTTGGCGAAGATGGA 2699
 DB 202 CCTCTTAGATGCTTCTCTTATGCTCAAGTGCATGCTGGGATGTGGAGAAAGATGCG 261

QY 2700 CTGTCTCTTATAGCAGGTGTTTATTAACCAAGATGTTAAGTGCAGAGAAATGTAT 2759
 DB 262 TTCTCTTAGTAATCAGATCTATTAATCACTAATGTTGAAATGCAAGCGAGAGATGTT 321

QY 2760 GATAAGATATCATCATGCTTCAGATTTGGTGCATCTTTAATGGATCCCAATAGTTCTTG 2819
 DB 322 GACAGGACATAGTATGCTTCAGACAGGTGCTCCATGATGGATCCAAACCACTTCCCTG 381

QY 2820 TTACTGGTACTTCAGAGGTATGAACTTCCGAGGCTTTTAAACAGACCATATCTACAA- 2878
 DB 382 ATGATCATGCTCAGCGCTTTGAACTCTATCAGCTCTTTCAGCAGCGCTGACTATGGGA 441

QY 2879 -----AGACCAGGATTTGATTAAACAATATAATACACTAATAGAA 2918
 DB 442 AGATTCAAGTTCTGAGGTTTACCATAAGGACGTCGTTACAGAGAAACAACTCTGATCGAA 501

QY 2919 GAAATGCTTCAGGTCTCATCTATATTTGGTGGTGGAGCTTATGTTACTTGGAGTGGGAAT 2978
 DB 502 GAGATGCTCTCCTCATCATCATGCTTGTGGGAGAAAGATTCAACCTCGGGTTGGACAG 561

QY 2979 GTGACCAAGAGAGGTGACATGAGAGAAATCATTTCACTTGTGTTGCAATGAAACCATG 3038
 DB 562 GTGGCTGCCACAGATGAAATCAAGAGGGAGATTATCCATCAGCTGAGCATCAAACTATG 621

QY 3039 CCACACAGTGCCATTGCCAAAATTTTACCTGAGATGAAATTAATGAAATGCTGCTTAGAG 3098
 DB 622 GCTCACAGCAGCTGTGGAAGTCTCTGCTCTGAAGATGAGAACAGAGACCGGCGTAGAG 681

QY 3099 AATGTCATAAACAAAGTGGCCACATTTAAGAAACCAAGGTGTATCAGGCCATGAGTTTAT 3158
 DB 682 AGCGTTCATCAGTCCGTAGCACATTTCAAGAAACCTGGGCTTCACAGGACGAGCATGTAT 741

QY 3159 GAACTAAAGATGAATCACTCTGAAGACTTCAATATGCTACTTTTATCATCTACTCCAAACC 3218
 DB 742 GAGCTGAGCCAGAGTGTGCCAAGAGTTCACCTGTATTTTATCATTTCTCCAGGGCA 801

QY 3219 CAGCATAGCAGGCTGAACATATGCAAGAAAGAGAGAAACCAAGAAACCAAGATGAA 3278
 DB 802 GAGCAGTCCAAAGCAGGAGAGCTCAGCGGAAATTTGAAAGAGAGAAATAAAGAGATACA 861

QY 3279 GCATTCGCCGCCACCAACCACTCTCTGAATTTCTGCCCTGCTTTCAGCAAGGTATTAACCTT 3338
 DB 862 GCATCTCCCTCTCCGGCTTTGCCACCGTTTCTGCCCTTTGTTCTGCGAGTCTGGTTAAACATC 921

QY 3339 CTCACACTGTGATATCATGATGTACATTTCTCAGGACCGTATTTTGAGCGGGCAATPAGACACA 3398
 DB 922 TTGCACTGTGAGCTCATGCTGTACATCATGCGGNAACGATCTTCGAGTGGGCTGTAGAGCAT 981

QY 3399 GATTCTAACTTGTGGACCGAAGGGATGCTCCAAATGGCTTTTTCATATTTCTGGCATTTGGGT 3458
 DB 982 CACGGGTCTGCTGGTTCAGAGTCCATGCTACAGAGGGTGTCTGCTTTGATCGGGATGGCT 1041

QY 3459 TTACTAGAGAGAGAGCAACAGCTTCAAAAAGCTCTCTGAAGCAAGAAGT-----AACATTTGAC 3515
 DB 1042 CTCACGAGAGAGAGCAACCACTTGGAGATGCGTGGAGGGCATGTGCGAGACCTTTCAAC 1101

QY 3516 TTTTATCATAGGCTTCAAGATTTGGGAGTTTCAG-----CCATGATATACAAATG 3566
 DB 1102 TTCAACAGAGATTTCAAAGCTGTGTATGACCAACATACTCCCGAGCATCTTAGCT 1161

QY 3567 CTTTGTGAAAAAATCTCAAAGAAATTTCCCAAGTTAGAGGCGCAGAGGACATGATAACGTGG 3626
 DB 1162 ATGCTGGAGACCTTGCAGAACCGCCCTCTCCGGAAGCCCAACAGGACATGATCAGGTGG 1221

QY 3627 ATACTTCAGATGTTTGACACAGTGAAGCGATTAAGAGAAAAATCTTGTGTTTAAATGTAGCA 3686
 DB 1222 TTGCTAAAGATGTTTAAATGCAATTAAGAAAGATAAGAGA-----GTGTTTCATCCAGCAGC 1275

QY 3687 ACCACATCAGGATCGGATCTATTAAGATGATGAGTACTCTCATGATTAAGAAAGCA 3746
 DB 1276 CCGGTGGCCGAGCGGAGGGAACCCATAATGGAGGAGAGCTCAAGAGACAGGAGCAAGCA 1335

QY 3747 GAAACGAAAAAGAAAGCTGAAGCTCTAGGCTTACATCGCCAGAGATCATGGCTCAGATG 3806
 DB 1336 GAGAGAAAAAGAAAGCCGAGATGCCAGACTGCGCGGAGAGAGATCATGSCCCAGATG 1395

QY 3807 TCTGCTTTACAGAAAACTTCATGAAACTCAATACTCATGTATGACAAATACATCAGAA 3866
 DB 1396 TCTGAGATGTCAGCGGCACTTCATTGACGAAAAACAAAGAGCTCTTCCAGCAGACCTTAGAG 1455

QY 3867 ATGCTCTGGGAAGAGATTTCCATTTAGGAGGAGGAGGAGCCAGCAGCTAGTGAATAC 3926
 DB 1456 CTGGACACCTCTGCTCTGCCACTCTT-----GACAGCAGCCCTCCCGCTTTTCAGACGCA 1509

QY 3927 TCTAGAAATGCTTTTGGGCTCTAAAGGGGTCCATCTGTACTGAAAAAGAGGTGCTGAGC 3986
 DB 1510 GCTCTTACAGCACTGGGCCCCAGCAGACAGACAGCTCCCTGTAACCGAGACAGTTTGTCAAC 1569

QY 3987 TGCATCTTTGCAAGAGAAACAGAGGTGAAATAATAGAAAAATATGCCCCATGTTATCG 4046
 DB 1570 TGTATATTATGTCAAGAGGAGCAAGAGGTGACTGTGGGAGCAGGCGGATGCTTGGCA 1629

QY 4047 GCCTGTGTCCAGAAATCTACTGCTTAACCCAGCAGCGGGGAAAAACCCATAGAACTCTCA 4106
 DB 1630 GCATTTGTCAGAGGTCAACCGTCTGTCTAAAGACAGGACGAGAAACCATTCGCGAGCCCA 1689

QY 4107 GGAGAGCCCTAGACCCACTTTTCATGATCCAGACTTGGCATATGGAATTTATACAGA 4166
 DB 1690 GAAAAA-----TATGATCCATTTATTCATGACCCCGATCTGTCTTGTGGGACACACATGCG 1746

QY 4167 AGCTGTGTGTCATGTAATGACGACAGTGTGCTGGCAGAAAGTATTTTGAAGCTGTAC----- 4221

Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.jp,
URL: <http://genome.gsc.riken.jp/>, Tel. 81-45-503-9222,
Fax: 81-45-503-9216)

COMMENT
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in Riken.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: <http://genome.gsc.riken.jp/>
URL: <http://phantom.gsc.riken.jp/>

FEATURES

source

1..3058
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM DB:6430402E23"
/db_xref="taxon:10090"
/clone="6430402E23"
/sex="male"
/tissue_type="olfactory brain"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
<1..691
/note="unnamed protein product; putative
ubiquitin protein ligase E3 component n-recognin 1
(MGDI|GI:1277977, GI|NM_009461, evidence: BLASTN, 100%,
match=2560)"
/codon_start=2
/protein_id="BAC37160.1"
/db_xref="GI:26347023"
/translation="GVAPPEELFANSAEPEPALSLSLPTNLFLLFQRYWDTIRPL
LQWCGDPAALSLKRSKSAVRYPRKNSLIHELPEYSLCLNOASHFRCPADDERK
HPVLCLFCAILCSQICQBIYNGEVGACVYFHALHCGAGVCI FLKRECRVILVEG
KARGCAYPAPYLDEYGETDPLKRGNPLHLSRERYKLHVMQHCIIIEIARSQETN
OMLFGFNWOLL"
3032..3037
/note="putative"
3058
/note="putative"

ORIGIN

Query Match 11.7%; Score 740.2; DB 3; Length 3058;
Best Local Similarity 82.4%; Pred. No. 8.9e-148;
Matches 904; Conservative 0; Mismatches 148; Indels 45; Gaps 3;
QY 5255 TGGGGTAACCTCCGCTGAGGAACATGACATACCAATCTCTGAGAGAGAGATACAGTGCAC 5314
DB 1 TGGAGTAGCTCCGCTGAGAACTGTTGGCAATCTGCTGAGAGAGAAATTCAGTGCAC 60
QY 5315 CTGTAGCTATCTATCTTTACCTACAAATTTGTTCTGCTCTCCAGGAATATTTGGGATAC 5374
DB 61 CTGTAGCTATCTATCTTTACCCACAAATTTGTTCTGCTCTTTCCAGGAATATTTGGGATAC 120
QY 5375 TGTAAGCCCTTCTCCAGAGGGTGTGACAGATCTGCTCTTACTTAACTGTTTGAAGCA 5434
DB 121 CATTAAGCCCTTCTACTAGAGGGTGTGAGAGATCTGCTCTTACTCAAGTCTTTGAAGCA 180
QY 5435 AAAAAACACCGTGTGAGTACCTAGAAAAGAAATAGTTTATGATAGCTTCTCTGATGA 5494
DB 181 GAAAGTGTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
QY 5495 CTATAGTGTCTCTGATCAATCAAGTCTCTATTTTCAAGTGTGCGGATCCAGTGTGATGATGA 5554
DB 241 CTACAGTGTCTCTTAAATCAGGCTTCTACCTTTAGTGTGCGGATCCAGTGTGATGATGA 300
QY 5555 GCGAAGATCTCTGCTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5614
DB 301 GCGAAGATCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTGCTGCT 360
QY 5615 CTGCGAGGAATTTGTGAACCGGGAAGAGGTGTGAGCTTGTGATTTTTCACGCACTTTCAC 5674
DB 361 TTGCCAGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420

QY 5675 TGGAGCCGAGTCTGCAATTTTCTTAAATATCAGAGAAATCCGAGTGGTCTGGTTGAAGG 5734
DB 421 TGGTCTCGAGTCTGCAATTTTCTTAAATATCAGAGAAATCCGAGTGGTCTGGTTGAAGG 480
QY 5735 TAAAGCCAGAGGCTGTGCTTATCCAGCTCTTACTTGGATGAATATGAGAAACAGACCC 5794
DB 481 AAAAGCCAGAGGCTGTGCTTATCCAGCTCTTACTTGGATGAATATGAGAAACAGACCC 540
QY 5795 TGGCTCTGAAGAGGGCAACCCCTTCAATTTATCTGAGGCGGTATCCGAGCTCAATTT 5854
DB 541 AGGGCTAAAGAGAGGAAACCCCACTTCAATTTATCTCGGAGCGGTATCCGAGCTCAATTT 600
QY 5855 GGTCTGGCAACCAACACTGCTTATATAGAGAGATTCCTAGGAGCCAGAGACTTAATCAGAT 5914
DB 601 GGTCTGGCAACAGCACTGCTTATATAGAGAGATTCCTAGGAGCCAGAGACTTAATCAGAT 660
QY 5915 GTTATTTGGATTTCAACTGGCAGTTACTCTGAGCTTCAACTCTGCTCAAGACAATCACA 5974
DB 661 GCTATTTGGATTTCAACTGGCAGTTACTCTGAGCTTCAAGACAATCACAATCATGA 720
QY 5975 ATGACGACAGTAGTAAGGCTGATTTCAAAATATATGAGAAACCTTCTGAGGCGTGGGAA 6034
DB 721 GTGACATCAATAATAAGACTGATCTAAATTTCTAGAGAACTTTCTGAGGAC-GGGGAG 779
QY 6035 TATTGGAGGCTCTTTTGTCTCCATGCTCCAGTTCCTTACATCAATAAATAATTTCTTAAT 6094
DB 780 TATTGGAGGCTCTTTTGTATCCATGCTCCAGTTCCTTACATCAATAAATAATTTCTTAAT 839
QY 6095 GGAGTATTGCTTTCAATTTAGCAACATATGCTTCTCAGGAAAAA-AGGACATAGATCAAT 6153
DB 840 GGAATATTGCTTTCAATTTATCAACATTAAGCTTTCAAGGGGAAAAA-CAAGACATAGATTAAT 899
QY 6154 CTGT-----TTTATGCTGCTAGT 6170
DB 900 GTTTATATGTTCTAGAACACTAAAGAAATGCTTTGTTTCATCCAGTGTCTATTTCTGCTAAT 959
QY 6171 ATTTCCAGGAATTTATTTCCCTTCAATAATTTGCTCTCATTTTATTTATTTATTTCACTTG 6230
DB 960 ATTTCCAGGAATTTATTTCCCTTCAATAATTTGCTCTCATTTTATTTATTTATTTCACTTG 1019
QY 6231 GTAGATAGTCACTCAACAGATTTGTAGACATTTTATGTTGGTTAACTTTCTTGTCAA 6290
DB 1020 GTTATAGGTCACATTAAGCAATTTGTGACATTTTCTCCATCTGCTTAACATCTCTGAC 1079
QY 6291 TTTTGTATTTGTTGTTT 6307
DB 1080 CTTTGTATTTGTTGTTT 1096
RESULT 9
AJ820333
LOCUS
DEFINITION
AJ820333 KN206 Bos sp. cDNA clone C0006018c05, mRNA linear EST 13-SRP-2004
ACCESSION
AJ820333
VERSION
AJ820333.1 GI:51887809
KEYWORDS
EST.
SOURCE
Bos sp.
ORGANISM
Bos sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi;
Mammalia; Rutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovinae; Bos.
REFERENCE
1 (bases 1 to 838)
McGuire, K. and Glass, E.J.
Unpublished ESTs, McGuire and Glass
Unpublished (2004)
CONTACT: McGuire K
Genomics and Genetics
Roslin Institute
Roslin, Midlothian, EH25 9PS, UNITED KINGDOM
Single pass sequencing. Bases called and trimmed with phred
v0.020425.c. Vector identified by crossmatch with the -minscore 20
and -minmatch 12 options. Vector: pBluescriptII(SK+) R. Site

Db	61	CATTGTTTTTCCACTATTACTTTGGGTAACCTCGGCTAGGAACTGCATACCAATTCG	120	AUTHORS	Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuina,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watanuki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
Qy	5293	CAGAAGGAGTACAGTGCACTCTGTAGCTATCTATCTTTACTACAAATTTGTCCTGC	5352	TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multipipillary sequencer
Db	121	CAGAAGGAGTACAGTGCACTCTGTAGCTATCTATCTTTACTGCAAAATTTGTCCTGC	180	JOURNAL MEDLINE	Genome Res. 10 (11), 1757-1771 (2000)
Qy	5353	TCCTCCAGCAATATTTGGGATCTGTAAAGCCCTTCTCCAGAGGTGGTGCAGATCCCTG	5412	PUBMED	20530913
Db	181	TCCTCCAGCAATATTTGGGATCTGTAAAGCCCTTCTCCAGAGGTGGTGCAGATCCCTG	240	REFERENCE	11076861
Qy	5413	CCTTACTAACTGTTTGAAGCAAAAACACCGTGTCTCAGGTACCCCTAGAAAAAGAAATA	5472	AUTHORS	4
Db	241	CCTTACTAACTGTTTGAAGCAAAAACACCGTGTCTCAGGTACCCCTAGAAAAAGAAATA	300	TITLE	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Qy	5473	GTTTGATAGAGCTTCTGTAGCTATAGTGTCTCTGTAATCAAGTTCCTATTTTCAGGT	5532	JOURNAL MEDLINE	Functional annotation of a full-length mouse cDNA collection
Db	301	GTTTGATAGAGCTTCTGTAGCTATAGTGTCTCTGTAATCAAGTTCCTATTTTCAGGT	360	REFERENCE	Nature 409, 685-690 (2001)
Qy	5533	GCCACGGTCTGCAGATGATGAGCGAAGCATCTCTCTGCTTCTCTGCGGCTA	5592	AUTHORS	5
Db	361	GCCACGGTCTGCAGATGATGAGCGAAGCATCTCTCTGCTTCTCTGCGGCTA	420	TITLE	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Qy	5593	TACTATGTTCTCAGACATTTGTCGCAGCAAAATTTGAAACGGGGAAGAGTTTGGAGCTT	5652	JOURNAL MEDLINE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Db	421	TACTATGTTCTCAGACATTTGTCGCAGCAAAATTTGAAACGGGGAAGAGTTTGGAGCTT	480	REFERENCE	Nature 420, 563-573 (2002)
Qy	5653	GCATTTTTCACGCATTTCTGTGGAGCGGAGTCTGCATTTTCTTAAAAATCAGAGAAT	5712	AUTHORS	6 (bases 1 to 3627)
Db	481	GCATTTTTCACGCATTTCTGTGGAGCGGAGTCTGCATTTTCTTAAAAATCAGAGAAT	540	TITLE	Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akanira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
Qy	5713	GCGAGTGTCTGTTGAAGTAAAGCCAGAGGCTGCTATCCAGCTCCCTTACTTCTGG	5772	JOURNAL MEDLINE	Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
Db	541	GCGAGTGTCTGTTGAAGTAAAGCCAGAGGCTGCTATCCAGCTCCCTTACTTCTGG	600	COMMENT	CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge. Please visit our web site for further details. URL:http://genome.gsc.riken.jp/
Qy	5773	ATGAATATGAGAAACAGACCTCGGCTCAAGAGGGGCAACCCCTTCATTTATCTCGTG	5832	FEATURES	Location/Qualifiers 1. .3627 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM DB:C920004H05" /db_xref="taxon:10090" /clone="C920004H05" /cell_type="thymic cells" /tissue_type="thymus" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="2 days neonate" 1044. .53626 /note="unnamed protein product; putative similar to (A PROTEIN SIMILAR TO UBIQUITIN-PROTEIN LIGASE E3 COMPONENT N-RECOGNIN (UBIQUITIN-PROTEIN LIGASE E3-ALPHA)) (FRAGMENT) [Homo sapiens] (SFTR Q9H578,
Db	601	ATGAATATGAGAAACAGACCTCGGCTCAAGAGGGGCAACCCCTTCATTTATCTCGTG	660	source	
Qy	5833	AGCGGTATCGGAGCTCCATTTGGTCTGGCAACAACATCGCATATAGAAGA	5884		
Db	661	AGCGGTATCGGAGCTCCATTTGGTCTGGCAACAACATCGCATATAGAAGA	712		
RESULT 11	AK083320	3627 bp	mRNA	linear	HTC 03-APR-2004
LOCUS	AK083320				
DEFINITION	Mus musculus 2 days neonate thymus cells cDNA, RIKEN full-length enriched library, clone:C920004H05 product:similar to (A PROTEIN SIMILAR TO UBIQUITIN-PROTEIN LIGASE E3 COMPONENT N-RECOGNIN (UBIQUITIN-PROTEIN LIGASE E3-ALPHA)) (FRAGMENT) [Homo sapiens], full insert sequence.				
ACCESSION	AK083320				
VERSION	AK083320.1	GI:26350448			
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus				
ORGANISM	Mus musculus (house mouse)				
REFERENCE	1 Carninci,P. and Hayashizaki,Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999)				
AUTHORS	Carninci,P. and Hayashizaki,Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	20499374				
PUBMED	11042159				
REFERENCE	3				

evidence: FASTY, 78.8%ID, 100%length, match=749")

```
/codon_start=1
/protein_id="BAC38864.1"
/db_xref="GI:26350449"
/translation="MFRYAVDILTWKESLPEDLEVBKSDTYCYCMLFNDEVHTYEQ
VYITLQKAVNCTQKBAIPATVDRDRSRVRYGDFQDOAKTVIVRNTSRQTKPLK
VOMHSSVAHONFGLKALSLWGLSGYDGLRILCOVLOEPPDGENSSLVDRMLK
NDSKLKMGARSVYHQLPMSLLMDLYKXKLPAIRPAKRYERLQDYVYTDHREPSVA
DLSVQIFTPVPSLARMLLITENLMTVIIKAPMDHLKRDQAGRFQFRIYALQAFKFR
VLSLILDLKVLISKTEWSEDLRFQFLOGFAPLELLKMQMGDITTRQVQGHIME
PEWEAAFTLQMLKTHVISMVQWALDEKVLIEAYKKLAVLTQCHGGFTDGEPTTL
SICGHSEVETIRYCVSOKSVIHLPIISRLLAGLHVLLSKSEVAKYPPELLPLSLEPPL
LIHLPLRCLVLAQVHWRNGFSLVNOIYYVHNKCRREMFDDIYMLQTVGSMM
DNHFLMILSRPELQLPSTPDYGRPSSEVTHKDVQVQNTLIEMLYLILMLVGE
RNFPGVQVATDEIRKELIHLQSLIKPMASSELVLSLPDENKTEGMSVIESVAHFR
KPLCTGRMYLPEKACGFNLYFIHFWAHEHGSKASEAQRKLKRENKBDTALPPPALP
PCPLFASLIVILQCDMLYIMGTILQWAVEHHSKAWSEMLQRLVHLIEMALQEBKH
HLENAVEGHVOTFTFKIKSPGDAPHNPSILAMLETIQNAPSLBAHKDMIRWLKLM
FNAIKKIRECSSSPVAEAEAGTIMBESSRDKDKAERKRKAEIARLRREKIMQMPFMO
RHPD"
```

ORIGIN

```
Query Match      10.9%; Score 687.8; DB 3; Length 3627;
Best Local Similarity 56.0%; Pred. No. 8.4e-171;
Matches 1472; Conservative 0; Mismatches 1107; Indels 51; Gaps 7;

QY 1245 GAAGAGTAAATGTCAGGCGAGAAATAATTTCTTCACTGATATAAATAATGTCGTAGAA 1304
DB 1005 GAAGATGTATGCCAGAACTTCAACATTTTGTATTATTTGCTATGTTTCGATATGAGTAGAT 1064
QY 1305 ATGACTATATGGAAGGAAAGAACTCCCTCTGNACTCCAGATAGGGAGAAAT 1364
DB 1065 ATACTGACCTGGGAAAGAAAGTGAATTCCTGGAAGACTTGAAGTGGCAGAGAGAGT 1124
QY 1365 GAAAGATATATTTGTGCTTTTCAATGATGAACACCATTCATATGACACAGCTATATAC 1424
DB 1125 GACACCTACTACTGCTGTTTATGATGAGGTTTCAACCTATGAGCAGTCAATTTAT 1184
QY 1425 AGCCTCAAAAGAGCTTTGACTGTGAGCTGCGAGAGCCAGTTGATACCACTGCCATT 1484
DB 1185 ACCCTTCAGAAAGCTGTGAACGTACACAGAGGAAGCCATTTGGCTTTGCAACTACAGTT 1244
QY 1485 GACAAAGAGGCTGTGGGCTGTAAAGCGGAGCTTATGCTGCTTCCAGGAGCAAG 1544
DB 1245 GATCGAGATGGCCGTAGTCTGTCCGATATGGAAGTTTCCAGTACTGTGATCAAGCAAG 1304
QY 1545 GAAGATATAAGAGTCAATTCAGAAATGTCTCTCAACATCCATTCATGTAGAGTATTA 1604
DB 1305 ACAGTCAATTTGAGGAACACCAGCAGACAGACC---AAGCGCTCAAGTTCAAGTTATG 1361
QY 1605 CACTCAGAGATTATGCTCATCAGAAATTTGCTTTGCTGTTGCTTGGATGAAACAA 1664
DB 1362 CACTCTCGTGGCTGCTCATCAGAAATTTTGGTTTGAAGCTCTGTGCTGGCTGGGAAGT 1421
QY 1665 ATTATGAGTATTCAGTGAATTTAGGAGATCTTTTGGCAGCATGCTTGGCAGCATGCCCTTAGAGAGAA 1724
DB 1422 GTTATTGGATCTCAGATGGCCCTTCAGAGATTTTGTGTCAAGTTGGATTTACAAAGAGT 1481
QY 1725 CCTGACTCGGAGAAATCCCTGTCTCATAGCAGGTAAATGCTTTGGGATGCAAGCTTTAT 1784
DB 1482 CCAGATGGCGAAACATCTCTCTGCTGCTGCGACAGACTGATGCTTAATGATTCCAATATGG 1541
QY 1785 AAAGGTGCCCGTAAGATCTCTCATGAATTTGATCTTCAGCAGATTTTTTTTATGGAGATGGAA 1844
DB 1542 AAAGGGCTAGGAGTGTGATACACAGTTGTTTCATAGCAGAGCTGCTCATGGACCTCAAG 1601
QY 1845 TACAAAAAATCTTTTGTATGGAATTTTGGAGATTTATATAACAACTGCGAGAAAGATAT 1904
DB 1602 TATAAGAGCTGTTTCGGCTTCGATTTGCTTAAACAACTATGAGCGTTTTCAGAGAGATAT 1661
QY 1905 ATCAGTGTATCATCAGCAGAGTATCTCTATAAATGCACTTTCAGTTTCAGATGTTTACT 1964
DB 1662 GTGACAGATGACCAACAGCAGAGAGTTTTCAGTGGCAGAGCTCTCGGTTTCAGATATTCA 1721
```

```
QY 1965 GTTCCTACTCTGCTCGACATCTTATTGAAGACGACAAATGTTATCTCTGTCTACTACTGAA 2024
DB 1722 GTTCCTTCTACTTTCGCGAAATGCTCTCTCACAGAGAGAACCTGATGACCGTTATCATTAAG 1781
QY 2025 ACTCTGCTAGAGTTTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2078
DB 1782 GCTTTCATGACCAATTTTGAACACAGAGATGCCAGGCGAGATTCAGTTTGAACCGTAC 1841
QY 2079 CAGGGTTATAGCAGGACAAATTTGGGAAGATATATGCAAGTATATATGTCACCTTAAAGTAT 2138
DB 1842 ACTGCCCTCAAGCTCTCAAGTTTCAGGAGAGTCCAGAGCTCATCTTAGATCTCAAGTAT 1901
QY 2139 ATCCTGATCAGCAAAACCCACATATGCAAGAAAGATTGAAGATGCAAGTTCCTTTGAAGGT 2198
DB 1902 GTATTGATAGCAAAACCAACGAGAGTGGTTCAGATGAGCTGAGGCGAGAAAGTTCTTTACAAGGG 1961
QY 2199 TTTGAGATCTTTTGAAGATTTCTTACTCTGATGCGGGAATGGAAGAAATCCGAAGACAG 2258
DB 1962 TTGATGCTCTCTTGGAAATTTCTGAAGTGCATGCGGGAATGGAACCGGATCAGCGCTCAG 2021
QY 2259 GTTGGCAACACATTTGAAGTGGATCTGATTTGGAGGCTGCCATTTGCTATATACAGATGCAA 2318
DB 2022 GTGGGACAGCACATTTGAGATGAGCCAGAGTGGAGAGCAGCTTTCACACTGCGAGATGAAG 2081
QY 2319 TTGAAGAAATATTTTACTCATGTTTCAAGAGTGGTGTCTTGTGATGAGAACTCTTTACTTT 2378
DB 2082 CTGACACAGCTCATCTCAATGTTGCGAGGACTGGTGTGCTCTGGACGAAAAAGTTTAAAT 2141
QY 2379 GTGGCTTATAAAGAAATGTCACAAAGCTGTGATGAGTGCAGTACCAAGTTTTCATATCTAGT 2438
DB 2142 GAAGCTTACAGAAATATGCTGGCTGTGTCGACACAGTGTCTGCGGGAATTTACTGATGGT 2201
QY 2439 AGCAAGACAGT---AGTACAATCGTGTGACATAGTTTGGAAAAACAAAGTCTTACAGAGTA 2495
DB 2202 GAACAGCCAAATCACACTCAGTATTTGTGACACTCGGTGGAAACCATCAGATCTGTGTT 2261
QY 2496 TCTGAGGATCTTTGTAAGCATACATCTGCCACTCTCTAGAGACCTTCTGCTGGTCTTCATGTA 2555
DB 2262 TCCCAAGAAAAAGTTAGCAATTCACCTCCCAATTTCTCGCTTGTCTGCTGAGGTTTGCATGTA 2321
QY 2556 CGTTTAAAGCAGGCTGGTCTCTCTTCAAGACTGCATGAATTTGTGTCTTTTGGAGGACTTT 2615
DB 2322 TTGTTAAGCAAAAGTGAAGTGGCATATAAATTTCCAGAGCTCTCTCTCTTAAGTGAACATG 2381
QY 2616 CAAGTAGAGTACTAGTGAATATCTTTTACGTTGTCTGTTGTGTTGTTGCTGCCAGGTTGTT 2675
DB 2382 AGCCCAACCATGTTGATAGAACTCTCTTAGATGTTCTTGTCTTATGCTCAAGTGCAT 2441
QY 2676 GCTGAGATGTGCGAAGAAATGGAAGTCTCTTTATAGCCAGGTGTTTATTAACCAAGAT 2735
DB 2442 GCTGGGATGTGGAAGAAATGGCTTCTCTTAGTAAATCAGATCTATTAACCAATAAT 2501
QY 2736 GTTAAAGTGCAGAGAAAGATGATGATGAAGATATCATCTGCTTCAGATTTGGTGCATCT 2795
DB 2502 GTGAAATGCGGCGAGAGATGTTTCGACAGGACATAGTATGCTTTCAGACAGGTGCTCTCC 2561
QY 2796 TTAATGGATCCCAATTAAGTTCTTGTGTTCTGTTGTTCTGAGGTATGAACTGCGCAGGCT 2855
DB 2562 ATGATGGACCCAAACCACTTCTCTGATGATCATGCTCAGCGGCTTTGNACTCTATCAGCTC 2621
QY 2856 TTTAAACAAGACATATCTACAAA-----AGACCAGGATTTGATTT 2894
DB 2622 TTCAGCAGCCCTGACTATGGGAAGAGATTCAGTTCTGAGGTTTACCCATAAGGACGCTGTT 2681
QY 2895 AAAACAATATAATACACTAATAGAGAAATGCTTTCAGGTTCTCTATCTATATTTGGGTGAG 2954
DB 2682 CAGCAGAAACCACTCTGATCGAAGAGATGCTCTACCTCATCATCTATGCTTGTGGAGAA 2741
QY 2955 CGTTATGTACCTGGAGTGGAAATGTGACAAAGAGAGGTTCACATATGAGAGAAATCAATT 3014
DB 2742 AGATTCAACCCCTGGGTTGGACAGGTTGGCTGCCACAGATGAATTCAGAGGGGAGATTATC 2801
```


Db 361 GGAATTGTGAACGGGAAGAGGTTGGAGCTTGCAATTTTTCAGCACTTCACTGTGGAGC 420
QY 5681 CGGAGTCTGCATTTTCTTAAATATCAGAGAAATGCCGAGTGGTCTCTGTTGAAGGTTAAAGC 5740
Db 421 CGGAGTCTGCATTTTCTTAAATATCAGAGAAATGCCGAGTGGTCTCTGTTGAAGGTTAAAGC 480
QY 5741 CAGAGGCTGTGCTATCCAGCTCTTACTTGGATGAATATGGAAGAACAGACCCCTGGCT 5800
Db 481 CAGAGGCTGTGCTATCCAGCTCTTACTTGGATGAATATGGAAGAACAGACCCCTGGCT 540
QY 5801 GAAGAGGGGCAACCCCTTCAATTTATCTGTGAGCGGTATCGAAGCTCATTGTTGCTG 5860
Db 541 GAAGAGGGGCAACCCCTTCAATTTATCTGTGAGCGGTATCGAAGCTCATTGTTGCTG 600
QY 5861 GCAACAACACTGCTATTAAGAGAGATTGCTAGGAGCCCAAGAGACTAATCAGATGTTATT 5920
Db 601 GCAACAACACTGCTATTAAGAGAGATTGCTAGGAGCCCAAGAGACTAATCAGATGTTATT 660
QY 5921 TGGATTCACTGGCAGTTACTG 5942
Db 661 GGGATTCACTGGCAGTTACTG 682

RESULT 13

BU155953
LOCUS
DEFINITION AGENCOURT_7966363 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:6165232
5', mRNA sequence.

ACCESSION BU155953

VERSION BU155953.1 GI:22669485

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 898)

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC/DCTD/BTP

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LIAW13523 row: j column: 17

High quality sequence stop: 606.

Location/Qualifiers

1..898

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:6165232"

/tissue type="melanotic melanoma"

/lab_host="DH10B (phage-resistant)"

/clone_lib="NIH MGC 72"

/note="Organ: skin; Vector: pCMV-SPORT6; Site: 1: NotI;

Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dr.

Average insert size 2 kb. Library constructed by Life

Technologies."

ORIGIN

Query Match 10.6%; Score 668.8; DB 5; Length 898;

Best Local Similarity 99.6%; Pred. No. 5.9e-166;

Matches 681; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 5626 TTGTGAACGGGAAGAGGTTGGAGCTTGCAATTTTTCAGCACTTCACTGTGGAGCGGAG 5685

Db 12 TTGTGAACGGGAAGAGGTTGGAGCTTGCAATTTTTCAGCACTTCACTGTGGAGCGGAG 71

QY 5686 TCTGCATTTTCTTAAATATCAGAGAAATGCCGAGTGGTCTCTGTTGAAGTAAAGCCAGAG 5745
Db 72 TCTGCATTTTCTTAAATATCAGAGAAATGCCGAGTGGTCTCTGTTGAAGTAAAGCCAGAG 131
QY 5746 GCTGTGCCCTATCATCTCTTACTTGGATGAATATGGAAGAACAGACCCCTGGCTGAAGA 5805
Db 132 GCTGTGCCCTATCATCTCTTACTTGGATGAATATGGAAGAACAGACCCCTGGCTGAAGA 191
QY 5806 GGGGCAACCCCTTCAATTTATCTGTGAGCGGTATCGGAAGCTCCTCAATTTGGTCTGCAAC 5865
Db 192 GGGGCAACCCCTTCAATTTATCTGTGAGCGGTATCGGAAGCTCCTCAATTTGGTCTGCAAC 251
QY 5866 AACACTGCATTTATAGAGAGATTGCTAGGAGCCCAAGAGACTAATCAGATGTTATTGGAT 5925
Db 252 AACACTGCATTTATAGAGAGATTGCTAGGAGCCCAAGAGACTAATCAGATGTTATTGGAT 311
QY 5926 TCAACTGGCAGTTACTGTGAGCTCCAACCTCTGAGCTCCAACCTCTGAGCTCCAACCTCAAGCAATCAAAATGACGACGT 5985
Db 312 TCAACTGGCAGTTACTGTGAGCTCCAACCTCTGAGCTCCAACCTCTGAGCTCCAACCTCAAGCAATCAAAATGACGACGT 371
QY 5986 AGTAAAGGCTGATTCAAAATTTATGGAATCTTCTGAGGCTGGGAAAGTATTGGAGGCT 6045
Db 372 AGTAAAGGCTGATTCAAAATTTATGGAATCTTCTGAGGCTGGGAAAGTATTGGAGGCT 431
QY 6046 CTTTGTCTCCATGTCAGGTTCACTTACATCAATTAATTTCTTAATGGAGTATTGCT 6105
Db 432 CTTTGTCTCCATGTCAGGTTCACTTACATCAATTAATTTCTTAATGGAGTATTGCT 491
QY 6106 TTCAATTAGCAAAATATGCTTTCAGAGGAAAGGACATAGATCAATCTGTTTTATG 6165
Db 492 TTCAATTAGCAAAATATGCTTTCAGAGGAAAGGACATAGATCAATCTGTTTTATG 551
QY 6166 CTAGTATTTCCAGGAATTTATTTCCCTTCATATTTTGTCTCATTTTCAATTTTATTCATCC 6225
Db 552 CTAGTATTTCCAGGAATTTATTTCCCTTCATATTTTGTCTCATTTTCAATTTTATTCATCC 611
QY 6226 ACTGCTAGATGAAGTCAGTCAGTCAAAAGTGTAGACATTTTATGTTT-GGTTAACTCTT 6284
Db 612 ACTGCTAGATGAAGTCAGTCAGTCAAAAGTGTAGACATTTTATGTTTGGTGGTAACTCTT 671
QY 6285 CTGCAATTTTGTATTTGGTGTGTTT 6308
Db 672 CTGCAATTTTGTATTTGGGTTTT 695

RESULT 14

BU18516/c

LOCUS

DEFINITION

UI-H-PH1-bfk-c-21-0-UI.s1 NCI CGAP_FH1 Homo sapiens cDNA clone

UI-H-PH1-bfk-c-21-0-UI 3', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

EST.

BU18516.1

GI:23284731

REFERENCE

1 (bases 1 to 668)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: James Martin

cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa

DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Clone distribution information can be obtained

from Dr. M. Bento Soares, bento-soares@uiowa.edu

Seq primer: M13 FORWARD

POLYA=Yes.

Location/Qualifiers

FEATURES

```
source
1. .668
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-FH1-bfk-c-21-0-UI"
/tissue_type="Cell Line"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP_FH1"
/note="Organ: Chondrosarcoma; Vector: p773-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP FH1 is a normalized cDNA library obtained from a cell line derived from grade I chondrosarcoma tissue. The library was constructed and normalized according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into p773-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is AGAATCCGCG. The cell line was provided by Dr. James Martin from the University of Iowa.
TAG_RISUSE=Human Chondrosarcoma Cell Line CS8 - Grade I
Chondrosarcoma
TAG_LIB=UI-H-FH1
TAG_SEQ=AGAATCCGCG"

ORIGIN
Query Match 10.5%; Score 664.8; DB 5; Length 668;
Best Local Similarity 99.7%; Pred. No. 6.2e-165;
Matches 666; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2607 GAGGACTTTCAAGTAGAGGTACTAGTGGAAATCTCTTTACGTGTCTGGTGTGGTTGCC 2666
DB 668 GAGGACTTTCAAGTAGAGGTACTAGTGGAAATCTCTTTACGTGTCTGGTGTGGTTGCC 609

QY 2667 CAGGTTGTCTGAGATGGCGAAGAAATGGATCTCTCTTTATGACCGAGGTGTTTTAT 2726
DB 608 CAGGTTGTCTGAGATGGCGAAGAAATGGATCTCTCTTTATGACCGAGGTGTTTTAT 549

QY 2727 TACCAGATGTTAAGTGCAGAGAAAGTATGATAAGATATCATCTTCAGATT 2786
DB 548 TACCAGATGTTAAGTGCAGAGAAAGTATGATAAGATATCATCTTCAGATT 489

QY 2787 GGTGCATCTTTAATGATCCCAATAGTCTCTTCTGTTACTGTTCTCAGAGGTATGAATT 2846
DB 488 GGTGCATCTTTAATGATCCCAATAGTCTCTTCTGTTACTGTTCTCAGAGGTATGAATT 429

QY 2847 GCCGAGGCTTTTAAAGACCATATCTACAAAGACGAGATTTGATTTAAACAATATAT 2906
DB 428 GCCGAGGCTTTTAAAGACCATATCTACAAAGACGAGATTTGATTTAAACAATATAT 369

QY 2907 ACATTAATAGAAATGCTTCAGGTCCTCATCTATATTTGGTTCAGGTTATGACCT 2966
DB 368 ACATTAATAGAAATGCTTCAGGTCCTCATCTATATTTGGTTCAGGTTATGACCT 309

QY 2967 GGAGTGGGAAATGTGACCAAGAGAGGTACCAATGAGAGAAATCACTTCACTTGTCTTTC 3026
DB 308 GGAGTGGGAAATGTGACCAAGAGAGGTACCAATGAGAGAAATCACTTCACTTGTCTTTC 249

QY 3027 ATTGAACCCATGCCACAGTGCATTTGCCAAAATTTTACCTGAGATGAAATATATGAA 3086
DB 248 ATTGAACCCATGCCACAGTGCATTTGCCAAAATTTTACCTGAGATGAAATATATGAA 189

QY 3087 ACTGGCTTACAGATGTCATAACAAAGTGGCCACATTTAAGAAACAGGTATCAGGC 3146
DB 188 ACTGGCTTACAGATGTCATAACAAAGTGGCCACATTTAAGAAACAGGTATCAGGC 129

QY 3147 CATGGAGTTTATGAACTTAAAGATGAATCACTGAAAGACTTCAATATGTACTTTTATCAT 3206
DB 121 CATGGAGTTTATGAACTTAAAGATGAATCACTGAAAGAGGTCAATGAGAGAAATCATTC 180
```

```
Db 128 CATGAGTTTATGAACCTAAAGATGAATCACTGAAAGACTTCAATATGTACTTTTATCAT 69
QY 3207 TACTCCAAACCCAGCATAGCAAGGCTGAACATATGCAAGAGAGAGAGAGAGAGAGAG 3266
Db 68 TACTCCAAACCCAGCATAGCAAGGCTGAACATATGCAAGAGAGAGAGAGAGAGAGAG 9
QY 3267 AACAAAGA 3274
Db 8 AACAAAAA 1

BG534574 756 bp mRNA linear EST 03-APR-2001
602553425F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4663182 5',
mRNA sequence.
BG534574
ACCESSION BG534574
VERSION BG534574.1 GI:13526116
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCW1465 row: a column: 07
High quality sequence stop: 751.
Location/Qualifiers
1. .756
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4663182"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_77"
/note="Organ: lung; Vector: pDNR-LJB (Clontech); Site 1:
SfiI (ggccctcgcc); Site 2: SfiI (ggccattggcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGGCCATTTATGCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGCGGCGCGGCGGCGCATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.9
kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."

ORIGIN
Query Match 10.4%; Score 656.6; DB 4; Length 756;
Best Local Similarity 95.9%; Pred. No. 9.9e-163;
Matches 730; Conservative 0; Mismatches 19; Indels 12; Gaps 5;

QY 2836 GGTATGAACCTTCCCGAGGCTTTTAAACAGACCATATCTACAAAGACGAGGATTTGATTA 2895
Db 1 GGTATGAACCTTCCCGAGGCTTTTAAACAGACCATATCTACAAAGACGAGGATTTGATTA 60

QY 2896 AACAAATAATACACTAATAGAGAAATGCTTCAGGTCCTCATCTATATTTGGGTGAGC 2955
Db 61 AACAAATAATACACTAATAGAGAAATGCTTCAGGTCCTCATCTATATTTGGGTGAGC 120

QY 2956 GTTATGTACTCGAGTGGGAAATGTGACCAAGAGAGAGGTCAATGAGAGAAATCATTC 3015
Db 121 GTTATGTACTCGAGTGGGAAATGTGACCAAGAGAGAGGTCAATGAGAGAAATCATTC 180
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2005, 07:46:33 ; Search time 130 Seconds
(without alignments)
5203.415 Million cell updates/sec

Title: US-10-758-672A-2
Perfect score: 9224
Sequence: 1 MADEAGGTERMEISAEPLQ.....ETARQETNQLFGFNWQLL 1749

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9224	100.0	1749	ADJ95442	Adj95442 Human Ubi
2	9224	100.0	1749	ADJ95442	ADJ95442 Human Ubi
3	9168	99.4	1738	ADJ95442	ADJ95442 Human Ubi
4	9087	98.5	1734	ADJ95442	ADJ95442 Human Ubi
5	9087	98.5	1734	ADJ95442	ADJ95442 Human Ubi
6	8617	93.4	1757	ADJ95442	ADJ95442 Human Ubi
7	8617	93.4	1757	ADJ95442	ADJ95442 Human Ubi
8	8617	93.4	1757	ADJ95442	ADJ95442 Human Ubi
9	8617	93.4	1757	ADJ95442	ADJ95442 Human Ubi
10	4768	51.7	908	ADJ95442	ADJ95442 Human Ubi
11	4327	46.9	1755	ADJ95442	ADJ95442 Human Ubi
12	4327	46.9	1755	ADJ95442	ADJ95442 Human Ubi
13	4278	46.4	1755	ADJ95442	ADJ95442 Human Ubi
14	4278	46.4	1755	ADJ95442	ADJ95442 Human Ubi
15	4209	45.6	811	ADJ95442	ADJ95442 Human Ubi
16	3457	37.5	1400	ADJ95442	ADJ95442 Human Ubi
17	3454	37.4	1400	ADJ95442	ADJ95442 Human Ubi
18	3160	34.3	1275	ADJ95442	ADJ95442 Human Ubi
19	3066	33.2	594	ADJ95442	ADJ95442 Human Ubi
20	2710.5	29.4	1109	ADJ95442	ADJ95442 Human Ubi
21	2278.5	24.7	1829	ADJ95442	ADJ95442 Human Ubi
22	2168	23.5	424	ADJ95442	ADJ95442 Human Ubi
23	1716	18.6	333	ADJ95442	ADJ95442 Human Ubi
24	1716	18.6	333	ADJ95442	ADJ95442 Human Ubi
25	1670	18.1	306	ADJ95442	ADJ95442 Human Ubi

26	1665	18.1	306	8	ADR09273	Human pro
27	1301.5	14.1	575	8	ADJ95442	Novel hum
28	1169.5	12.7	452	5	ABJ0168	Human pol
29	903.5	9.8	343	8	ADJ95442	Human pro
30	877	9.5	487	4	ABJ05917	Novel hum
31	790	8.6	258	4	ABJ05917	Novel hum
32	606.5	6.6	1225	5	ABJ05917	Novel hum
33	468	5.1	147	7	ABJ05917	Novel hum
34	453.5	4.9	741	8	ABJ05917	Novel hum
35	453.5	4.9	951	6	ABJ05917	Novel hum
36	449	4.9	389	4	ABJ05917	Novel hum
37	420	4.6	1456	4	ABJ05917	Novel hum
38	414.5	4.5	203	5	ABJ05917	Novel hum
39	404.5	4.4	709	8	ABJ05917	Novel hum
40	401	4.3	71	8	ABJ05917	Novel hum
41	360	3.9	153	4	ABJ05917	Novel hum
42	341	3.7	550	8	ABJ05917	Novel hum
43	311	3.4	783	4	ABJ05917	Novel hum
44	311	3.4	783	8	ABJ05917	Novel hum
45	302	3.3	437	8	ADN00361	Novel hum

ALIGNMENTS

RESULT 1
ADJ95442

ID ADJ95442 standard; protein; 1749 AA.

AC ADJ95442;

DT 03-JUN-2004 (first entry)

DE Human Ubiquitin ligase E3alpha I.

KW Human; enzyme; ubiquitin ligase; E3alpha I; ubiquitin-proteasome pathway;
KW Gene therapy; vaccine; muscular atrophy; cachexia; catabolic disorders;
KW cancer cachexia; renal cachexia; inflammatory cachexia;
KW muscle wasting disorder; metabolic acidosis; uremia; burn;
KW hyperthyroidism; Cushing's syndrome; fasting; denervation atrophy;
KW diabetes mellitus; sepsis; AIDS wasting syndrome; SNP;
KW single nucleotide polymorphism.

OS Homo sapiens.

Key Location/Qualifiers

FT Misc-difference 1568 /note= "May be Arg as the result of a single nucleotide polymorphism"

FT US6706505-B1.

PD 16-MAR-2004.

PF 28-NOV-2000; 2000US-00724126.

PR 08-MAR-2000; 2000US-0187911P.

PA (AMGE-) AMGEN INC.

PI Han H, Kwak K;

DR WPI; 2004-236723/22.

XX N-PSDB; ADJ95441.

XX New nucleic acid molecule, useful for preparing a composition for
diagnosing, treating or preventing diseases associated with human
E3alpha polypeptide, e.g., muscle atrophy.

PS Claim 1; SEQ ID NO 2; 104pp; English.

XX The invention relates to a new isolated nucleic acid molecule appearing
as ADJ95441 (or its complement) encoding a ubiquitin ligase E3alpha I

CC protein appearing as ADJ95442. Also included are a vector comprising the
CC nucleic acid, a host cell comprising the vector, a process of producing a
CC E3alpha I ubiquitin ligase polypeptide, a composition comprising the
CC nucleic acid molecule, a reagent comprising a detectably labelled
CC nucleotide, and a method for determining the presence of a human E3alpha
CC I ubiquitin ligase nucleic acid in a biological sample. The nucleic acid
CC molecule is useful for preparing a composition for diagnosing, treating
CC or preventing diseases associated with human E3alpha I polypeptide, e.g.
CC muscle atrophy, cachexia, catabolic disorders, cancer cachexia, renal
CC cachexia, inflammatory cachexia, muscle wasting disorders associated with
CC metabolic acidosis, uremia, burns, hyperthyroidism, Cushing's syndrome,
CC fasting, denervation atrophy, diabetes mellitus, sepsis and AIDS wasting
CC syndrome. The present sequence represents human E3alpha I.
XX
SQ

Sequence 1749 AA;

Query Match 100.0%; Score 9224; DB 8; Length 1749;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1749; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MADEEAGGTERMEISAELEPQTPORLASWDDQVDFYTAFLHLAQIVPEIYFAEMDPDLE 60

DB 1 MADEEAGGTERMEISAELEPQTPORLASWDDQVDFYTAFLHLAQIVPEIYFAEMDPDLE 60

QY 61 KOBESVQMSIFTPLEWYLFGEODPDICLEKLKHSFALQGRVFKSGETTYSCRDCAIDPT 120

DB 61 KOBESVQMSIFTPLEWYLFGEODPDICLEKLKHSFALQGRVFKSGETTYSCRDCAIDPT 120

QY 121 CVLQMDCFQDSVHKHRYKMHSTGGGFCDCGTEAWKTGPFVNHPEGRAGTIKENSRC 180

DB 121 CVLQMDCFQDSVHKHRYKMHSTGGGFCDCGTEAWKTGPFVNHPEGRAGTIKENSRC 180

QY 181 PLNEEVIQARKIPPSVIKTVEMTIWEEKELPELOIREKNERYCYLVFDEHSHYD 240

DB 181 PLNEEVIQARKIPPSVIKTVEMTIWEEKELPELOIREKNERYCYLVFDEHSHYD 240

QY 241 VYISLQALDCELAQAHLHTAIDKEGRRVAVAGAAQEAQEDIKSHSENVSOHPLHV 300

DB 241 VYISLQALDCELAQAHLHTAIDKEGRRVAVAGAAQEAQEDIKSHSENVSOHPLHV 300

QY 301 EVLHSEIMAHQFALRGSGMKNKIMSYSDFRQIFCOACLRBEPDSENCLISRLMLMDA 360

DB 301 EVLHSEIMAHQFALRGSGMKNKIMSYSDFRQIFCOACLRBEPDSENCLISRLMLMDA 360

QY 361 KLYKGARKILHELIFFSPFMEMEYKCLFAMEFYKYLQKEYISDDHRSISITALSVO 420

DB 361 KLYKGARKILHELIFFSPFMEMEYKCLFAMEFYKYLQKEYISDDHRSISITALSVO 420

QY 421 MFTVPTLARHLIEQNVISVITETLLEVLPEYLDNRNKNFQGYSDQKLGRRVAVICDLK 480

DB 421 MFTVPTLARHLIEQNVISVITETLLEVLPEYLDNRNKNFQGYSDQKLGRRVAVICDLK 480

QY 481 YLISKPTIWTERRMQFLGPRSFILKILTCMQGMEIEIRQVGOHIEVDPDWEAAIAIQW 540

DB 481 YLISKPTIWTERRMQFLGPRSFILKILTCMQGMEIEIRQVGOHIEVDPDWEAAIAIQW 540

QY 541 QLNKILLMFOEWACDBELLVAYKECHKAVMRCSTSFSSKTVVQSCGHSLETYSYV 600

DB 541 QLNKILLMFOEWACDBELLVAYKECHKAVMRCSTSFSSKTVVQSCGHSLETYSYV 600

QY 601 SEDLVSIHLPLSRTLAGLVRLSRGAVSRKLFHFVDFQVRLVEYPLRCLVLVAQV 660

DB 601 SEDLVSIHLPLSRTLAGLVRLSRGAVSRKLFHFVDFQVRLVEYPLRCLVLVAQV 660

QY 661 AEMWRRNGLSISQVFFYQDVKCREEMYDKDIIIMLQIGASLMDPNKFLLLVLQRYELAEA 720

DB 661 AEMWRRNGLSISQVFFYQDVKCREEMYDKDIIIMLQIGASLMDPNKFLLLVLQRYELAEA 720

QY 721 FNKTIKTDQDLIKQNTLIEEMQLVLIYVGRYVPGVGNVTKEEVTWREIHLICIEP 780

DB 721 FNKTIKTDQDLIKQNTLIEEMQLVLIYVGRYVPGVGNVTKEEVTWREIHLICIEP 780

QY 781 MPHSAIAKNLPENENNETGLENVINKVATPKPGVSGHGVYELKDESLKDFNMYFYHSK 840

DB 781 MPHSAIAKNLPENENNETGLENVINKVATPKPGVSGHGVYELKDESLKDFNMYFYHSK 840

QY 841 TOHSAEAMOKKRRKQENKDEALPPPPPEPCFAPSKVINLLNCDDIMMYILRTVPERAID 900

DB 841 TOHSAEAMOKKRRKQENKDEALPPPPPEPCFAPSKVINLLNCDDIMMYILRTVPERAID 900

QY 901 TDSNLTWTEGMLQMAFHIALGALLEKQOLQKAPBEEVTFDFVHKASRLGSSAMNTQMLLE 960

DB 901 TDSNLTWTEGMLQMAFHIALGALLEKQOLQKAPBEEVTFDFVHKASRLGSSAMNTQMLLE 960

QY 961 KLKGIPLQEGQKDMITWILQMPDFTVKRLREKSLCLIVATTSGSEIKNDEIHTHKEKABRK 1020

DB 961 KLKGIPLQEGQKDMITWILQMPDFTVKRLREKSLCLIVATTSGSEIKNDEIHTHKEKABRK 1020

QY 1021 RKAEEAARLHROKIMQAQMSALQKNFIETHKLMYDNTSEMPGKEDSIMESESTPAVSDYSRI 1080

DB 1021 RKAEEAARLHROKIMQAQMSALQKNFIETHKLMYDNTSEMPGKEDSIMESESTPAVSDYSRI 1080

QY 1081 ALGPKRGPSVTEKEVLTCTILCOBEOEVKIENNAMVLSACVOKSTALTQHRGKPIELSGEA 1140

DB 1081 ALGPKRGPSVTEKEVLTCTILCOBEOEVKIENNAMVLSACVOKSTALTQHRGKPIELSGEA 1140

QY 1141 LDPLPMDPLAYGTYTSGCHVMAHVCWKYFEAVQLSSQRIHVLDLPDLSGSEYLCPLC 1200

DB 1141 LDPLPMDPLAYGTYTSGCHVMAHVCWKYFEAVQLSSQRIHVLDLPDLSGSEYLCPLC 1200

QY 1201 KSLCNTVPIIPLQPKINSENADALAOULLTARMIQTVLARISYNTIRHAKENPIPIF 1260

DB 1201 KSLCNTVPIIPLQPKINSENADALAOULLTARMIQTVLARISYNTIRHAKENPIPIF 1260

QY 1261 FNOQMGDSTLRFHSITLSFGVSESSIKYSNSIKEMVTLFATTIYRIGLKVPPDERDRVPM 1320

DB 1261 FNOQMGDSTLRFHSITLSFGVSESSIKYSNSIKEMVTLFATTIYRIGLKVPPDERDRVPM 1320

QY 1321 TWSTCAFTIQAIENTLLGDEGKPLFCALQNRQHNGLKALMQFAVAQRITCPOVLIQKHLVR 1380

DB 1321 TWSTCAFTIQAIENTLLGDEGKPLFCALQNRQHNGLKALMQFAVAQRITCPOVLIQKHLVR 1380

QY 1381 LLSVVLPMIKSEDTCLLSIDLPHVLVGAFLAPPSLYWDDPVDLPQSSVSSSYNHLVLP 1440

DB 1381 LLSVVLPMIKSEDTCLLSIDLPHVLVGAFLAPPSLYWDDPVDLPQSSVSSSYNHLVLP 1440

QY 1441 LITMAHMLQIILLTVDTGLPLAQVQSDSEAHSASSFPAEISQYTSIGSDIPGWYLVWS 1500

DB 1441 LITMAHMLQIILLTVDTGLPLAQVQSDSEAHSASSFPAEISQYTSIGSDIPGWYLVWS 1500

QY 1501 LKNGITPYLRCALPFHYLLGVTTPPEELHTNSABGEYSALCSYLSLPTNLFLLPQBYWDT 1560

DB 1501 LKNGITPYLRCALPFHYLLGVTTPPEELHTNSABGEYSALCSYLSLPTNLFLLPQBYWDT 1560

QY 1561 VRPILQRCADPALLNCLKQNTVYRPRKNSLIELPDDYSLCLNOASHFCRPSADDE 1620

DB 1561 VRPILQRCADPALLNCLKQNTVYRPRKNSLIELPDDYSLCLNOASHFCRPSADDE 1620

QY 1621 RKHPVLCLFCGAILCSQNICCOEIVNGEVEGACIFALHCGAGVCIFLKIRCRVVLVVG 1680

DB 1621 RKHPVLCLFCGAILCSQNICCOEIVNGEVEGACIFALHCGAGVCIFLKIRCRVVLVVG 1680

QY 1681 KARGCAYPAPYLDEYGETDPGLKRGNPILHLSRERYKHLVWQOHCIIIEIARSQETNOM 1740

DB 1681 KARGCAYPAPYLDEYGETDPGLKRGNPILHLSRERYKHLVWQOHCIIIEIARSQETNOM 1740

QY 1741 LFGFNWQLL 1749

DB 1741 LFGFNWQLL 1749

RESULT 2

ADS86864

ID ADS86864 standard; protein; 1749 AA.

XX

AC ADS86864;

XX 16-DEC-2004 (first entry)
 XX Human E3alpha ubiquitin ligase, hE3alpha1 protein sequence.
 XX Human; E3alpha ubiquitin ligase; huE3alpha1;
 KW ubiquitin-proteasome pathway; chromosome 15 q; rapid muscle wasting;
 KW fasting; metabolic acidosis; muscle degeneration; kidney failure;
 KW renal cachexia; uremia; diabetes mellitus; sepsis; AIDS wasting syndrome;
 KW cancer cachexia; negative nitrogen balance; burn; Cushing's syndrome;
 KW inflammatory cachexia; hyperthyroidism; denervation atrophy;
 KW protein/tissue wasting; energy-protein malnutrition; muscle atrophy;
 KW gene therapy; enzyme.
 XX Homo sapiens.
 OS US2004185037-A1.
 XX 23-SEP-2004.
 XX 15-JAN-2004; 2004US-00758672.
 XX 08-MAR-2000; 2000US-0187911P.
 PR 28-NOV-2000; 2000US-00724126.
 XX (HANH/) HAN H.
 PA (KWAK/) KWAK K.
 XX Han H, Kwak K;
 PI WPI; 2004-707854/69.
 DR N-PSDB; ADS86863.
 XX Novel isolated human E3alpha ubiquitin ligase nucleic acid molecule
 PT useful for treating and/or preventing renal cachexia or inflammatory
 PT cachexia.
 XX Claim 13; SEQ ID NO 2; 115pp; English.
 CC The present invention relates to new orthologue of human E3alpha
 CC ubiquitin ligase, huE3alpha1 and huE3alpha1I. Most intracellular proteins
 CC are degraded through the ubiquitin-proteasome pathway. Proteins are
 CC marked for proteasomal degradation by conjugation of ubiquitin to the
 CC protein. Conjugation of the ubiquitin molecule involves the activation by
 CC E1 enzyme, subsequent transfer to E2 enzyme, which serves as a carrier,
 CC and then interacts with a specific E3 ligase family member. E3 ligase
 CC binds to proteins targeted for degradation and catalyses the transfer of
 CC ubiquitin from the E2 carrier enzyme to the target protein. E3 ligase
 CC determines the specificity of the system. The E3alpha family is the main
 CC family of intracellular ligases and is involved in the N-end rule pathway
 CC of protein degradation. E3alpha enzyme binds directly to the primary
 CC destabilising N-terminal amino acid and catalyses ubiquitin conjugation
 CC thereby targeting the protein for degradation. The human E3alpha gene is
 CC located on chromosome 15 q. Increased proteolysis through the ubiquitin-
 CC proteasome pathway has been determined to be a major cause of rapid
 CC muscle wasting including, fasting, metabolic acidosis, muscle
 CC degeneration, kidney failure, renal cachexia, uremia, diabetes mellitus,
 CC sepsis, AIDS wasting syndrome, cancer cachexia, negative nitrogen
 CC balance, burns, Cushing's syndrome, inflammatory cachexia,
 CC hyperthyroidism, denervation atrophy, protein/tissue wasting, and energy-
 CC protein malnutrition. E3alpha plays a role in the overall increase in
 CC ubiquitination that is associated with and may mediate muscle atrophy in
 CC catabolic and other disease states. Treatment may be administered by gene
 CC therapy, cell therapy and antisense therapy methods. The present sequence
 CC is human E3alpha1 protein sequence.
 XX Sequence 1749 AA;
 SQ Query Match 100.0%; Score 9224; DB 8; Length 1749;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1749; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 MADEBAGGTERMEISAEPLQTPQRLASWDDQVDFYTAFLHHLAQVLVPEIYFAEMDPDLE 60

1 MADEBAGGTERMEISAEPLQTPQRLASWDDQVDFYTAFLHHLAQVLVPEIYFAEMDPDLE 60
 61 KQEEVSQMSIFTPLEWYLFGEHPDICLEKLKHSAGFQLCGRVFKSGTETYSRCDCAIDPT 120
 61 KQEEVSQMSIFTPLEWYLFGEHPDICLEKLKHSAGFQLCGRVFKSGTETYSRCDCAIDPT 120
 121 CVLQMDPCFQDSVHKHRYKQHTSTGGGFCDCGDTBAWTKTGPFCVNHPEGRAGTIKENSRC 180
 121 CVLQMDPCFQDSVHKHRYKQHTSTGGGFCDCGDTBAWTKTGPFCVNHPEGRAGTIKENSRC 180
 181 PLNEEVIQVARKIPPSVIVVVEMTIWEKEKLPELQIREKNERVYCYLVNDEHSHYDH 240
 181 PLNEEVIQVARKIPPSVIVVVEMTIWEKEKLPELQIREKNERVYCYLVNDEHSHYDH 240
 241 VIYSQRALDCELAQAQLHTTAIDKEGRRVAVAGAYAAQCEAKEDIKSHSENVSQPLHV 300
 241 VIYSQRALDCELAQAQLHTTAIDKEGRRVAVAGAYAAQCEAKEDIKSHSENVSQPLHV 300
 301 EYLHSEIMAHQKFAIRLGSWMNKIMSYSDPQIIFCQACLRREPDSNPCLISRLMLWDA 360
 301 EYLHSEIMAHQKFAIRLGSWMNKIMSYSDPQIIFCQACLRREPDSNPCLISRLMLWDA 360
 361 KLYKGARKILHILIFSSPFMEYKYLPAFMEFYKYKLOKEYISDDHRSISITALSVO 420
 361 KLYKGARKILHILIFSSPFMEYKYLPAFMEFYKYKLOKEYISDDHRSISITALSVO 420
 421 METVPTLARHLIEEQNVISVITETLLEVLPEYLDNRNKNFQYSDKLRVYAVICDLK 480
 421 METVPTLARHLIEEQNVISVITETLLEVLPEYLDNRNKNFQYSDKLRVYAVICDLK 480
 481 YILISKPTIWTERRMQFLGFRSFLKILTCMQGMEIRRQVQGHIEVDPDEAAIAIQM 540
 481 YILISKPTIWTERRMQFLGFRSFLKILTCMQGMEIRRQVQGHIEVDPDEAAIAIQM 540
 541 QLNKILLMPQEWACDEBELLVAYKECHKAVNRCTSPFISSSKTVVQSCGHSLETYSRV 600
 541 QLNKILLMPQEWACDEBELLVAYKECHKAVNRCTSPFISSSKTVVQSCGHSLETYSRV 600
 601 SEDLVSIHLPLSRTLAGLHVRLSGAVSRLEHFVSFEDQVEVLVEYVLRCLVLVAQVV 660
 601 SEDLVSIHLPLSRTLAGLHVRLSGAVSRLEHFVSFEDQVEVLVEYVLRCLVLVAQVV 660
 661 AEMWRNGLSLISQVFFYQDVKCREMYDKDIIIMLQIGASLMDPNKFLLLVLQRYELASA 720
 661 AEMWRNGLSLISQVFFYQDVKCREMYDKDIIIMLQIGASLMDPNKFLLLVLQRYELASA 720
 721 FNKTIISTKQDQDLIKQYNTLIEBMLQVLIYIGERVYVGVNTKEEVTWREIHLICIEP 780
 721 FNKTIISTKQDQDLIKQYNTLIEBMLQVLIYIGERVYVGVNTKEEVTWREIHLICIEP 780
 781 MPHSAIAKNLPENNENNETGLENVINKVATFKKPGVSGHGVYELKDESLKDFNMYFVHYSK 840
 781 MPHSAIAKNLPENNENNETGLENVINKVATFKKPGVSGHGVYELKDESLKDFNMYFVHYSK 840
 841 TQHSKAEHMQKRRQKQENKDEALPPPPPEFCFAPSKVINLNCIDIMYILATVFERAID 900
 841 TQHSKAEHMQKRRQKQENKDEALPPPPPEFCFAPSKVINLNCIDIMYILATVFERAID 900
 901 TDSNLTWTEGMLQMAFHIALGLLEBKQLOKAPBEVTFDFVHKASRLGSSAMNTQMLLE 960
 901 TDSNLTWTEGMLQMAFHIALGLLEBKQLOKAPBEVTFDFVHKASRLGSSAMNTQMLLE 960
 961 KLKGIPLQEGQDMITWILQMPDVTVKRLREKSLVATTSSESITKNDIETHDKAEK 1020
 961 KLKGIPLQEGQDMITWILQMPDVTVKRLREKSLVATTSSESITKNDIETHDKAEK 1020
 1021 RKAEEARLHRQKIMAQMSALQKNFIETHKLMYDNTSEMPGKEDSIWESESTPAVDYSRI 1080
 1021 RKAEEARLHRQKIMAQMSALQKNFIETHKLMYDNTSEMPGKEDSIWESESTPAVDYSRI 1080
 1081 ALGPRGSGVTEKEVLTCTILCOEBOEVKIENNAMVLSACVQKSTALTQHRGKPIELSGA 1140

1081 ALGPKRGPSTVEKVTLCILQEQEVKLENNAMVLISACVQKSTALTOHRGKPIELSGEA 1140
1141 LDPLFMDPDLAYGTTGSCGHVNHAVCWQKYPEAVQLSSQQRHIVDLFDLESGEYLCPLC 1200
1141 LDPLFMDPDLAYGTTGSCGHVNHAVCWQKYPEAVQLSSQQRHIVDLFDLESGEYLCPLC 1200
1201 KSLCNTVPIIPILOPKINSEADALAOQLTLARWITQVLARI SGVNI RHAKGENPIPIF 1260
1201 KSLCNTVPIIPILOPKINSEADALAOQLTLARWITQVLARI SGVNI RHAKGENPIPIF 1260
1261 FNQMGDSTLEPHSILSFVSESSIKYSNISIKEMVILFATTTIYRIGLKVPVDERPRVPM 1320
1261 FNQMGDSTLEPHSILSFVSESSIKYSNISIKEMVILFATTTIYRIGLKVPVDERPRVPM 1320
1321 TWSTCAFTQAIENLGDGKPLFGALQNRQNGIKALMQFAVAQRITCPQVLIOKHLVR 1380
1321 TWSTCAFTQAIENLGDGKPLFGALQNRQNGIKALMQFAVAQRITCPQVLIOKHLVR 1380
1381 LLSVVLPNIKSEDTPCLLSIDLPHVLVGVAVLAPPSLYWDDPVDLOPSSVSSSYNHLXLFH 1440
1381 LLSVVLPNIKSEDTPCLLSIDLPHVLVGVAVLAPPSLYWDDPVDLOPSSVSSSYNHLXLFH 1440
1441 LITMAHMLQILITVTGTLPLAQVQEDSEAHGASSFFAEISQYTSIGSGCDIPGWLWVS 1500
1441 LITMAHMLQILITVTGTLPLAQVQEDSEAHGASSFFAEISQYTSIGSGCDIPGWLWVS 1500
1501 LKNGITPYLRCAALPFHYLLGVTTPPELHTNSAGEYSALCSYLSLPTNLFLFOEYWD 1560
1501 LKNGITPYLRCAALPFHYLLGVTTPPELHTNSAGEYSALCSYLSLPTNLFLFOEYWD 1560
1561 VRPLLQRCADPALLNCLKQKNTVVRYPKRNLSLIELDPDYSCLLNQASHFRCPASDDE 1620
1561 VRPLLQRCADPALLNCLKQKNTVVRYPKRNLSLIELDPDYSCLLNQASHFRCPASDDE 1620
1621 RKHPVLCLFCGAILCSQNICCOBIVNGEVGACIFHALHCGAGVCIFLKIRECRVVLV 1680
1621 RKHPVLCLFCGAILCSQNICCOBIVNGEVGACIFHALHCGAGVCIFLKIRECRVVLV 1680
1681 KARGCAYPAPYLDEYGETDPLKRGNPLHLSRERYKHLVWQOHCIIIEIARSQETNOM 1740
1681 KARGCAYPAPYLDEYGETDPLKRGNPLHLSRERYKHLVWQOHCIIIEIARSQETNOM 1740
1741 LFGFNWQLL 1749
1741 LFGFNWQLL 1749

RESULT 3
ABP58330
ID ABP58330 standard; protein; 1738 AA.

XX AC ABP58330;
XX DT 07-APR-2003 (first entry)
XX DE Human cell growth, differentiation and death protein CGDD-1.
XX KW CGDD-1; cell growth; cell differentiation; cell death; human; cytostatic;
KW antiatherosclerotic; hepatotropic; antiinflammatory; antipsoriatic;
KW antianemic; ophthalmological; auditory; anticonvulsant;
KW cerebroprotective; neutropic; neuroprotective; anti-HIV; antiallergic;
KW neuroleptic; tranquilizer; immunosuppressive; anti-parkinsonian;
KW antiasthmatic; antithyroid; antidiabetic; dermatological; nephrotropic;
KW antirheumatic; antiarthritic; antidiabetic; vulnular; virucide;
KW antibacterial; fungicide; antiparasitic; protozoacide; antihelminthic;
KW antifertility; gynaecological; ubiquitin protein ligase; enzyme;
KW gene therapy.
XX OS Homo sapiens.
XX PN W0200297032-A2.
XX XX
XX PD 05-DEC-2002.

XX 05-APR-2002; 2002WO-US011152.
XX 06-APR-2001; 2001US-0282110P.
PR 11-APR-2001; 2001US-0283294P.
PR 26-APR-2001; 2001US-0286820P.
PR 27-APR-2001; 2001US-0287228P.
PR 16-MAY-2001; 2001US-0291662P.
PR 18-MAY-2001; 2001US-0291846P.
PR 25-MAY-2001; 2001US-0293727P.
PR 01-JUN-2001; 2001US-0295263P.
PR 01-JUN-2001; 2001US-0295340P.
PR 15-JAN-2002; 2002US-0349705P.
XX (INCY-) INCYTE GENOMICS INC.
XX Azimzai Y, Au-Young JK, Batra S, Baughn MR, Becha SD,
PI Borowsky ML, Burford N, Ding L, Elliott VS, Emerling BM, Gandhi AR;
PI Gietzen KJ, Griffin JA, Hafalia AJA, Honchell CD, Lal PG, Lee SY;
PI Lu DAM, Arvizu CS, Ramkumar J, Reddy R, Sanjanwala MM, Tang YT;
PI Walia NK, Wang YE, Warren BA, Xu Y, Yang J, Yao MG, Yue H;
PI Zebbarjadian Y;
XX WPI; 2003-140453/13.
DR N-PSDB; ABZ24689.
XX Novel human proteins associated with cell growth, differentiation and
PT death, useful for treating, diagnosing or preventing cancer,
PT developmental, neurological, reproductive or autoimmune/inflammatory
PT disorders.
XX Claim 1; Page 183-187; 238pp; English.
XX The present sequence is the protein sequence of human CGDD-1, a novel
CC protein associated with cell growth, differentiation and death. The
CC sequence is predicted from Incyte clone 1351608CB1, which was isolated
CC from a paragonomic tumour tissue cDNA library. Structural features
CC establish the protein as being associated with cell growth,
CC differentiation and death, with further evidence suggesting it to be a
CC ubiquitin protein ligase. The invention is based on novel human CGDD-1 to
CC -21 proteins (see ABP58330-50), the polynucleotides encoding them (see
CC ABZ24689-709), and to the use of these for the diagnosis, treatment or
CC prevention of cell proliferative disorders including cancer,
CC developmental disorders, neurological disorders, autoimmune disorders,
CC reproductive disorders, and disorders of the placenta, and in the
CC assessment of the effects of exogenous compounds on the activity and
CC expression of proteins and nucleic acids associated with cell growth,
CC differentiation and death
XX SQ Sequence 1738 AA;
Query Match 99.4%; Score 9168; DB 6; Length 1738;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1738; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 12 MEISAEPLQTPORLASWMDQVDFTAFHLHQAQLVPELYPAEMDPDLKQESVQMSIF 71
DB 1 MEISAEPLQTPORLASWMDQVDFTAFHLHQAQLVPELYPAEMDPDLKQESVQMSIF 60
QY 72 TPLEWYLFGEHPDILCLEKLKHSAGFQLCGRVPKSGETTYSCRDCAIDPTCVLCMDCFQDS 131
DB 61 TPLEWYLFGEHPDILCLEKLKHSAGFQLCGRVPKSGETTYSCRDCAIDPTCVLCMDCFQDS 120
QY 132 VHKHRYKMHSTTGGGFCDCGTEAWKTGPPFCVNHPEGRAGTIKNSRCPLNEEIVQAR 191
DB 121 VHKHRYKMHSTTGGGFCDCGTEAWKTGPPFCVNHPEGRAGTIKNSRCPLNEEIVQAR 180
QY 192 KIFPSPVIVYVEMTWEEKELPPELQIREKNERYCYLVFNDEHHSYDHVYISLQALDC 251
DB 181 KIFPSPVIVYVEMTWEEKELPPELQIREKNERYCYLVFNDEHHSYDHVYISLQALDC 240
QY 252 ELAEEAQLHTTATDKEGRRRAVKAGAACQAEKEDIKSHSENVSOHPLHVEVLHSEIMAHQ 311

Db 241 ELAEAOQLHTTAIDKEGRAVAKAGAYAAQCEAKEDIKSHSENVSOHPLHVEVLHSEIMAHQ 300
Qy 312 KPALRLGSMWNKIMSYSSDFRQIFCOACLRBEPDSNPCLISRLMLWDAKLYKGARKILH 371
Db 301 KPALRLGSMWNKIMSYSSDFRQIFCOACLRBEPDSNPCLISRLMLWDAKLYKGARKILH 360
Qy 372 ELIFSSFFMEMEYKLLPAMEFVKYKQLOKEVYISDDHDSISITALSVMQFTVPTLARHL 431
Db 361 ELIFSSFFMEMEYKLLPAMEFVKYKQLOKEVYISDDHDSISITALSVMQFTVPTLARHL 420
Qy 432 IBEQNVISVITETLLEVLPEYLDNRNKNFQGYSDQKLAGRVAVICDLKYILISKPTIWT 491
Db 421 IBEQNVISVITETLLEVLPEYLDNRNKNFQGYSDQKLAGRVAVICDLKYILISKPTIWT 480
Qy 492 ERLRMQFLGFRSFLKILTCMQGMBEIRQVQGHIEVDPDWEAAIAIQMLKNILLMFOE 551
Db 481 ERLRMQFLGFRSFLKILTCMQGMBEIRQVQGHIEVDPDWEAAIAIQMLKNILLMFOE 540
Qy 552 WCACDEELLVAYKECHKAVMRCTSPISSTKVQSCGHSLETYSYRVEDLVSJHLPL 611
Db 541 WCACDEELLVAYKECHKAVMRCTSPISSTKVQSCGHSLETYSYRVEDLVSJHLPL 600
Qy 612 SRTLAGLHVRLSRLGAVSRILHEFVSPEDQVEVLVEYPLRCLVLAQVVAEMWRNGLSL 671
Db 601 SRTLAGLHVRLSRLGAVSRILHEFVSPEDQVEVLVEYPLRCLVLAQVVAEMWRNGLSL 660
Qy 672 ISQVFYQDVKCREEMDKDIIIMQIGASLMDPNKFLLLVLQRYELAEAFNKTISTKQD 731
Db 661 ISQVFYQDVKCREEMDKDIIIMQIGASLMDPNKFLLLVLQRYELAEAFNKTISTKQD 720
Qy 732 LIKQYNTLIEMLQVLIYVGERYPGVGNVTKKEVTWREIILHLLCIEPMPSIAKTNLP 791
Db 721 LIKQYNTLIEMLQVLIYVGERYPGVGNVTKKEVTWREIILHLLCIEPMPSIAKTNLP 780
Qy 792 ENNETETGLENVINKVATPKKGVSGHYVELKDBSLKDPNMFYHYSTQHSKAEHMOK 851
Db 781 ENNETETGLENVINKVATPKKGVSGHYVELKDBSLKDPNMFYHYSTQHSKAEHMOK 840
Qy 852 KRRKQENKDEALPPPPPPFPAPSKVINLLNCDIMMYILRTVFERAIDTDSNLMTEGML 911
Db 841 KRRKQENKDEALPPPPPPFPAPSKVINLLNCDIMMYILRTVFERAIDTDSNLMTEGML 900
Qy 912 QNAFHILALGLLEEKQLOKAPBEVTPFYHKASRLGSSANNIOMLLEKLGKIGIPOLEGQ 971
Db 901 QNAFHILALGLLEEKQLOKAPBEVTPFYHKASRLGSSANNIOMLLEKLGKIGIPOLEGQ 960
Qy 972 KDMITWILQMFDTVKRLREKSLIVATTSSGSIKNDEITHDKBAERKRKAARLHRQ 1031
Db 961 KDMITWILQMFDTVKRLREKSLIVATTSSGSIKNDEITHDKBAERKRKAARLHRQ 1020
Qy 1032 KIMQMSALQKQNFIEHKLMDNTSEMPGKEDSIMEEESTPAVSDYSRIALGPGRGPSVT 1091
Db 1021 KIMQMSALQKQNFIEHKLMDNTSEMPGKEDSIMEEESTPAVSDYSRIALGPGRGPSVT 1080
Qy 1092 EKEVLTCLCOBEQEVKIEENAMVLSACVQKSTALTQHRGKPIELSGEALDPLFMDPDIA 1151
Db 1081 EKEVLTCLCOBEQEVKIEENAMVLSACVQKSTALTQHRGKPIELSGEALDPLFMDPDIA 1140
Qy 1152 YGTYTGSCGHVHVCWQKYFEAVOLSSOORLHVLDLFDLESCEYLCPCKSCNTVPIPI 1211
Db 1141 YGTYTGSCGHVHVCWQKYFEAVOLSSOORLHVLDLFDLESCEYLCPCKSCNTVPIPI 1200
Qy 1212 PLQPKINSADALAQLLTLARWTQTLARISGVNIRHAKGNPIPIFNGQMGDSTLE 1271
Db 1201 PLQPKINSADALAQLLTLARWTQTLARISGVNIRHAKGNPIPIFNGQMGDSTLE 1260
Qy 1272 PHSILSPGVESIKYSNSIKEMVILPATTIYRIGLKVPDPDRPRVPMLTWSTCAFTIOA 1331
Db 1261 PHSILSPGVESIKYSNSIKEMVILPATTIYRIGLKVPDPDRPRVPMLTWSTCAFTIOA 1320
Qy 1332 IENILGDEGKPLFGALONRQHNGLKALMQFAVQAITCQVLIQKHLVRLLSVLPNKS 1391
Db 1321 IENILGDEGKPLFGALONRQHNGLKALMQFAVQAITCQVLIQKHLVRLLSVLPNKS 1380

RESULT 4

ADJ95459

ID ADJ95459 standard; protein; 1734 AA.

XX AC ADJ95459;

XX DT 03-JUN-2004 (first entry)

XX Human Ubiquitin ligase E3alpha I variant.

XX Human; enzyme; ubiquitin ligase; E3alpha I; ubiquitin-proteasome pathway;
gene therapy; vaccine; muscular atrophy; cachexia; catabolic disorders;
cancer cachexia; renal cachexia; inflammatory cachexia;
muscle wasting disorder; metabolic acidosis; uremia; burn;
hyperthyroidism; Cushing's syndrome; fasting; denervation atrophy;
diabetes mellitus; sepsis; AIDS wasting syndrome; SNP;
single nucleotide polymorphism.

XX Homo sapiens.

OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 225..246 /note= "Encoded by nucleotides 658-678 of SEQ ID 18 "

FT Misc-difference 1573 /note= "May be Tip as the result of a single nucleotide

FT polymorphism"

FT Misc-difference 1665..1666 /note= "Encoded by nucleotides 4933-4998 of SEQ ID 18 "

XX US6706505-B1.

XX 16-MAR-2004.

XX 28-NOV-2000; 2000US-00724126.

XX 08-MAR-2000; 2000US-0187911P.

XX (AMGE-) AMGEN INC.

XX Han H, Kwak K;

XX WPI; 2004-236723/22.

XX N-PSDB; ADJ95458.

XX New nucleic acid molecule, useful for preparing a composition for
diagnosing, treating or preventing diseases associated with human

PT	E3approximately polypeptide, e.g., muscle atrophy.
XX	Claim 19; SEQ ID NO 19; 104pp; English.
XX	The invention relates to a new isolated nucleic acid molecule appearing as ADJ95441(or its complement) encoding a ubiquitin ligase E3alpha I protein appearing as ADJ95442. Also included are a vector comprising the nucleic acid, a host cell comprising the vector, a process of producing the E3alpha I ubiquitin ligase polypeptide, a composition comprising the nucleic acid molecule, a reagent comprising a detectably labelled nucleotide, and a method for determining the presence of a human E3alpha I ubiquitin ligase nucleic acid in a biological sample. The nucleic acid molecule is useful for preparing a composition for diagnosing, treating or preventing diseases associated with human E3alpha I polypeptide, e.g. muscle atrophy, cachexia, catabolic disorders, cancer cachexia, renal cachexia, inflammatory cachexia, muscle wasting disorders associated with metabolic acidosis, uremia, burns, hyperthyroidism, Cushing's syndrome, fasting, denervation atrophy, diabetes mellitus, sepsis and AIDS wasting syndrome. The present sequence represents a human E3alpha I variant protein encoded by a cDNA assembled from previously isolated cDNA fragments.
XX	Sequence 1734 AA;
SQ	Query Match 98.5%; Score 9087; DB 8; Length 1734; Best Local Similarity 98.8%; Pred. No. 0; Matches 1728; Conservative 0; Mismatches 1; Indels 20; Gaps 1;
QY	1 MADEAGGTERMEISAEPLQTPQRLASWMDQVDFYTAFLHHLAQIVPEIYPAEMDPPDLE 60
DB	6 MADEAGGTERMEISAEPLQTPQRLASWMDQVDFYTAFLHHLAQIVPEIYPAEMDPPDLE 65
QY	61 KQESVQMSIFTPLEWYLFGEPPDICLEKXHSAPQLCGRVFKSGGETTYSRCDCAIDPT 120
DB	66 KQESVQMSIFTPLEWYLFGEPPDICLEKXHSAPQLCGRVFKSGGETTYSRCDCAIDPT 125
QY	121 CVLCHDCFDQSVHKNHRYKMHSTGGGFCDCGDTAWKTGFCVNHPEGRAGTIKENSRC 180
DB	126 CVLCHDCFDQSVHKNHRYKMHSTGGGFCDCGDTAWKTGFCVNHPEGRAGTIKENSRC 185
QY	181 PLNEEVIVQARKIFPSVIKYVEMTWIWEBEKELPPELQIREKNERYCYVLFNDEHSHYD 240
DB	186 PLNEEVIVQARKIFPSVIKYVEMTWIWEBEKELPPELQIREKNERYCYVLFNDEHSHYD 245
QY	241 VYISLQALDCELAELAAQLHTTAIDKEGRRAVAGAYAAQCEAKEDIKSHSENVSHPLHV 300
DB	246 VYISLQALDCELAELAAQLHTTAIDKEGRRAVAGAYAAQCEAKEDIKSHSENVSHPLHV 305
QY	301 EVLHSEIMAHQFALRLGSMWNKIMSYSDDFROIQCOACLRBEPDSENPCLISRLMLWDA 360
DB	306 EVLHSEIMAHQFALRLGSMWNKIMSYSDDFROIQCOACLRBEPDSENPCLISRLMLWDA 365
QY	361 KLYKGARKILHELIFSSFPFMEYKYLKAFMEFVKYKQKEVYISDDHDSISITALSQ 420
DB	366 KLYKGARKILHELIFSSFPFMEYKYLKAFMEFVKYKQKEVYISDDHDSISITALSQ 425
QY	421 MFTVPTLARHLIEBQNVISVITETLLLEVLPYLDNRNKNFQGYSDKLGRVAVICDLK 480
DB	426 MFTVPTLARHLIEBQNVISVITETLLLEVLPYLDNRNKNFQGYSDKLGRVAVICDLK 485
QY	481 YILISKPTIWTFLRMQFLGFRSFLKILTCMQGMEIRROVGOHLEVPDMEAAIAIQM 540
DB	486 YILISKPTIWTFLRMQFLGFRSFLKILTCMQGMEIRROVGOHLEVPDMEAAIAIQM 545
QY	541 QLKNILLMPQEWACDBELLILVAYKECHAVRCSTSFISSTKVVSQSGHSLKTSYRV 600
DB	546 QLKNILLMPQEWACDBELLILVAYKECHAVRCSTSFISSTKVVSQSGHSLKTSYRV 605
QY	601 SEDLVSIHPLSLSTLGLVRLSRLGAVSRHLHFVSFDFQVEVLVEYPLRCLVLVAQVV 660
DB	606 SEDLVSIHPLSLSTLGLVRLSRLGAVSRHLHFVSFDFQVEVLVEYPLRCLVLVAQVV 665
QY	661 AEWRRNGLSLISQVFFYQDVKCREMYDKDIIMLQIGASLMDPNKFLLLQRYELAEA 720

DB	666	AEWRRNGLSLSQVFFYQDVKCREMYDKDIIMLQIGASLMDPNKFLLLVLQRYELAEA 725
QY	721	FNKISTKQDQLIKQYNTLIEBMLQVLIYIVGERVPGVGNVTKEBVTMRILHLLCTEP 780
DB	726	FNKISTKQDQLIKQYNTLIEBMLQVLIYIVGERVPGVGNVTKEBVTMRILHLLCTEP 785
QY	781	MPHSAIAKNLPENNENNETGLNVINKVATPKKPGVSGHGVYELKDESUKDFNMYPHYSK 840
DB	786	MPHSAIAKNLPENNENNETGLNVINKVATPKKPGVSGHGVYELKDESUKDFNMYPHYSK 845
QY	841	TQHSKAEHQKRRKQENKDBALPPPPPPFCPAFSKVINLLNCDIMMYILRTVFERAID 900
DB	846	TQHSKAEHQKRRKQENKDBALPPPPPPFCPAFSKVINLLNCDIMMYILRTVFERAID 905
QY	901	TDSNLWTEGMLQMAFHIALALGLEEKQOLQKAPBEVTFDFYHKASRLGSSAMNLOMLE 960
DB	906	TDSNLWTEGMLQMAFHIALALGLEEKQOLQKAPBEVTFDFYHKASRLGSSAMNLOMLE 965
QY	961	KLKGIPOLEGOKDMITWILQMPDVTVRUREKSLIVATTSGSESIKNDEITHDKKABRK 1020
DB	966	KLKGIPOLEGOKDMITWILQMPDVTVRUREKSLIVATTSGSESIKNDEITHDKKABRK 1025
QY	1021	RKAEARLHRQKIMQMSALQKNFIETHKLMYDNTSEMPGKEDSIMBESPTAVSDYSRI 1080
DB	1026	RKAEARLHRQKIMQMSALQKNFIETHKLMYDNTSEMPGKEDSIMBESPTAVSDYSRI 1085
QY	1081	ALGPKRGSVTEKEVLTCTILCOEQEVKIEENAMVLSACVOKSTALTTOHRKGPIELSSEA 1140
DB	1086	ALGPKRGSVTEKEVLTCTILCOEQEVKIEENAMVLSACVOKSTALTTOHRKGPIELSSEA 1145
QY	1141	LDPLFMDPDLAYGTYTSCGHVMAVCWQKFEAVQLSSQQRHIVDLPLDESGEVLCPLC 1200
DB	1146	LDPLFMDPDLAYGTYTSCGHVMAVCWQKFEAVQLSSQQRHIVDLPLDESGEVLCPLC 1205
QY	1201	KSLCNTVPIPIPLQPKINSADALAQLLTLARMIQTVLARISGYNIRHAKENPPIPIF 1260
DB	1206	KSLCNTVPIPIPLQPKINSADALAQLLTLARMIQTVLARISGYNIRHAKENPPIPIF 1265
QY	1261	FNQMGDSTLBFHSLTSPGVESSIKYSNIIKEMVILPATYIRIGLKVPPDBRDRVPML 1320
DB	1266	FNQMGDSTLBFHSLTSPGVESSIKYSNIIKEMVILPATYIRIGLKVPPDBRDRVPML 1325
QY	1321	TWSTCAFTIQAENLILGDEGKPLFCALQNRQHGLKALMQFAVQIRITCPOVLIQKHLVR 1380
DB	1326	TWSTCAFTIQAENLILGDEGKPLFCALQNRQHGLKALMQFAVQIRITCPOVLIQKHLVR 1385
QY	1381	LILSVVLPNIKSEDTPCLLSIDLHFVLVGAVLAPPSLYWDDPVDLPQSSVSSSYNHLYLPH 1440
DB	1386	LILSVVLPNIKSEDTPCLLSIDLHFVLVGAVLAPPSLYWDDPVDLPQSSVSSSYNHLYLPH 1445
QY	1441	LITMAHMLQILLTVDTGTPLAQVQDSBEAHSASSFFAIEISQYTSGSGICDIPGMYLWVS 1500
DB	1446	LITMAHMLQILLTVDTGTPLAQVQDSBEAHSASSFFAIEISQYTSGSGICDIPGMYLWVS 1505
QY	1501	LKNGITPVLRCALPFHYLLGVTPEELHTNSAGEYSALCSYLSLPTNLFLLFQBYWDT 1560
DB	1506	LKNGITPVLRCALPFHYLLGVTPEELHTNSAGEYSALCSYLSLPTNLFLLFQBYWDT 1565
QY	1561	VRPLLQRCADPALNCLKQKNVTYVYPRKRNLSLIELPDDYSCLLNQASHFRCPSADDE 1620
DB	1566	VRPLLQRCADPALNCLKQKNVTYVYPRKRNLSLIELPDDYSCLLNQASHFRCPSADDE 1625
QY	1621	RKHPVLCIFCGAILCSQNICQOEIVNGBEVGCIFHALHCGAGVCIFLKIREFRVVLVEG 1680
DB	1626	RKHPVLCIFCGAILCSQNICQOEIVNGBEVGCIFHALHCGAGVCIFLKIREFRVVLVEG 1685
QY	1681	KARGCAYAPYLDYGETDPGLKRCNPLHLSPERVRKHLVWQOHCIIIEIARSQETNQ 1740
DB	1686	KARGCAYAPYLDYGETDPGLKRCNPLHLSPERVRKHLVWQOHCIIIEIARSQETNQ 1745
QY	1741	LFGFNQQL 1749

Db	1726	LFGENWQLL 1734	
RESULT 5			
ADN86881			
ID	ADN86881	standard; protein; 1734 AA.	
XX	AC	ADN86881;	
XX	DT	16-DEC-2004 (first entry)	
XX	DE	Variant human E3alpha1 ubiquitin ligase.	
XX	KW	Human; E3alpha ubiquitin ligase; huE3alpha1;	
XX	KW	ubiquitin-proteasome pathway; chromosome 15 q; rapid muscle wasting;	
XX	KW	fasting; metabolic acidosis; muscle degeneration; kidney failure;	
XX	KW	renal cachexia; uremia; diabetes mellitus; sepsis; AIDS wasting syndrome;	
XX	KW	cancer cachexia; negative nitrogen balance; burn; Cushing's syndrome;	
XX	KW	inflammatory cachexia; hyperthyroidism; denervation atrophy;	
XX	KW	protein/tissue wasting; energy-protein malnutrition; muscle atrophy;	
XX	KW	gene therapy; enzyme.	
XX	OS	Homo sapiens.	
XX	PN	US2004185037-A1.	
XX	PD	23-SEP-2004.	
XX	PF	15-JAN-2004; 2004US-00758672.	
XX	PR	08-MAR-2000; 2000US-0187911P.	
XX	PR	28-NOV-2000; 2000US-00724126.	
XX	PA	(HANH/) HAN H.	
XX	PA	(KWAK/) KWAK K.	
XX	PI	Han H, Kwak K;	
XX	DR	WPI; 2004-707854/69.	
XX	PT	Novel isolated human E3alpha ubiquitin ligase nucleic acid molecule	
XX	PT	useful for treating and/or preventing renal cachexia or inflammatory	
XX	PT	cachexia.	
XX	PS	Example 9; SEQ ID NO 19; 115pp; English.	
XX	CC	The present invention relates to new orthologues of human E3alpha	
XX	CC	ubiquitin ligase, huE3alpha1 and huE3alpha11. Most intracellular proteins	
XX	CC	are degraded through the ubiquitin-proteasome pathway. Proteins are	
XX	CC	marked for proteasomal degradation by conjugation of ubiquitin to the	
XX	CC	protein. Conjugation of the ubiquitin molecule involves the activation by	
XX	CC	E1 enzyme, subsequent transfer to E2 enzyme, which serves as a carrier,	
XX	CC	and then interacts with a specific E3 ligase family member. E3 ligase	
XX	CC	binds to proteins targeted for degradation and catalyses the transfer of	
XX	CC	ubiquitin from the E2 carrier enzyme to the target protein. E3 ligase	
XX	CC	determines the specificity of the system. The E3alpha family is the main	
XX	CC	family of intracellular ligases and is involved in the N-end rule pathway	
XX	CC	of protein degradation. E3alpha enzyme binds directly to the primary	
XX	CC	destabilising N-terminal amino acid and catalyses ubiquitin conjugation	
XX	CC	thereby targeting the protein for degradation. The human E3alpha gene is	
XX	CC	located on chromosome 15 q. Increased proteolysis through the ubiquitin-	
XX	CC	proteasome pathway has been determined to be a major cause of rapid	
XX	CC	muscle wasting including, fasting, metabolic acidosis, muscle	
XX	CC	degeneration, kidney failure, renal cachexia, uremia, diabetes mellitus,	
XX	CC	sepsis, AIDS wasting syndrome, cancer cachexia, negative nitrogen	
XX	CC	balance, burns, Cushing's syndrome, inflammatory cachexia,	
XX	CC	hyperthyroidism, denervation atrophy, protein/tissue wasting, and energy-	
XX	CC	protein malnutrition. E3alpha plays a role in the overall increase in	
XX	CC	ubiquitination that is associated with and may mediate muscle atrophy in	
XX	CC	catabolic and other disease states. Treatment may be administered by gene	
XX	CC	therapy, cell therapy and antisense therapy methods. The present sequence	
XX	CC	is a variant of human E3alpha1 ubiquitin ligase. The patent describes	
XX	CC	this sequence as containing a Trp at position 1568 as a result of a SNP	

CC	in the nucleotide sequence. However, the sequence contains a Pro at			
CC	position 1568.			
XX	Sequence	1734 AA;		
SQ	Query Match	98.5%; Score 9087; DB 8; Length 1734;		
	Best Local Similarity	98.8%; Pred. No. 0;		
	Matches 1728; Conservative	0; Mismatches 1; Indels 20; Gaps 1;		
QY	1	MADDEAGGTERMEISAELEPQTQRLASWMDQVDFYTAFLHLAQLVPEIYFAEMDPDLR 60		
DB	6	MADDEAGGTERMEISAELEPQTQRLASWMDQVDFYTAFLHLAQLVPEIYFAEMDPDLR 65		
QY	61	KQERSVQMSIPTLEWYLFGEEDPDIKLEKLKUSGAFQLCGRVFKGETTYSRCDCAIDPT 120		
DB	66	KQERSVQMSIPTLEWYLFGEEDPDIKLEKLKUSGAFQLCGRVFKGETTYSRCDCAIDPT 125		
QY	121	CVLGMDCTQDSVHKHRYKMTSTGGFCDCGDTBAWTKGPPCVNHEPGRAGTICKENSR 180		
DB	126	CVLGMDCTQDSVHKHRYKMTSTGGFCDCGDTBAWTKGPPCVNHEPGRAGTICKENSR 185		
QY	181	PLNEEVIYQARKIPPSVIKYVEMTIWBEKELEPPELQIREKNERYCYVLFNDEHHSYDH 240		
DB	186	PLNEEVIYQARKIPPSVIKYVEMTIWBEKELEPPELQIREKNERYCYVLFNDEHHSYDH 245		
QY	241	VIYSLQALDCELAQAQLHTTAIDKEGRRVAVAGAYAAQCAKEEDIKHSERNVSQHLHV 300		
DB	246	VIYSLQALDCELAQAQLHTTAIDKEGRRVAVAGAYAAQCAKEEDIKHSERNVSQHLHV 305		
QY	301	EVLHSEINAHOKPALRLGSMWNKIMSYSDDPROIFCOACLRPEEPSENPCILSRMLWDA 360		
DB	306	EVLHSEINAHOKPALRLGSMWNKIMSYSDDPROIFCOACLRPEEPSENPCILSRMLWDA 365		
QY	361	KLYGARKILHELIFSSFFMEYKYLKLFAMEPVKYYKQLQKEYISDDHRSISITALS 420		
DB	366	KLYGARKILHELIFSSFFMEYKYLKLFAMEPVKYYKQLQKEYISDDHRSISITALS 425		
QY	421	MFTVPTLARHLIEBONVISVITETLLVLPVLDNNKPNFGYSGDKLGRVAVICDLK 480		
DB	426	MFTVPTLARHLIEBONVISVITETLLVLPVLDNNKPNFGYSGDKLGRVAVICDLK 485		
QY	481	YLISKPTIWTERRMQFLGFRSPLKILTCQGMEEIRROVGOHIEVDPPDEAAIAIQM 540		
DB	486	YLISKPTIWTERRMQFLGFRSPLKILTCQGMEEIRROVGOHIEVDPPDEAAIAIQM 545		
QY	541	QLKNILLMFQWCACDELLLVAYKECHAVNRCSTSPISSTKVQSCGHSLETYSYRV 600		
DB	546	QLKNILLMFQWCACDELLLVAYKECHAVNRCSTSPISSTKVQSCGHSLETYSYRV 605		
QY	601	SEDLVSIHLPLSRTLAGLHVLRLSRLGAVSRLLHEFVSPEDQVEVLVYPLRCLVLVAQV 660		
DB	606	SEDLVSIHLPLSRTLAGLHVLRLSRLGAVSRLLHEFVSPEDQVEVLVYPLRCLVLVAQV 665		
QY	661	AMWRRNGLSLISQVFFYQDYKCEEMVDKDIIMLQIGASLMDPNKFLLLVLQRYELAE 720		
DB	666	AMWRRNGLSLISQVFFYQDYKCEEMVDKDIIMLQIGASLMDPNKFLLLVLQRYELAE 725		
QY	721	FNKTIQDQLIKQYNTLIEEMQLVLIYVGERVPGVGNVYKEVTMRREIHLICTEP 780		
DB	726	FNKTIQDQLIKQYNTLIEEMQLVLIYVGERVPGVGNVYKEVTMRREIHLICTEP 785		
QY	781	MPHSAIAKNLPENENNETGLENVINKVATFKKPGVSGHGVYELKDESLKDFNMYPHYSK 840		
DB	786	MPHSAIAKNLPENENNETGLENVINKVATFKKPGVSGHGVYELKDESLKDFNMYPHYSK 845		
QY	841	TQSKAEHQKRRKQENKDALPPPPPPPCFAPSKVINLLNCDIMMYILRTVFERAID 900		
DB	846	TQSKAEHQKRRKQENKDALPPPPPPPCFAPSKVINLLNCDIMMYILRTVFERAID 905		
QY	901	TDSNLTWTEGMLQMAFHILALGLLREKQOLQKAPBEVTFDFYHKASRLGSSAMNQMLLE 960		
DB	906	TDSNLTWTEGMLQMAFHILALGLLREKQOLQKAPBEVTFDFYHKASRLGSSAMNQMLLE 965		

QY 961 KLKGIPOLEGQDMITWILQMFDTVKRLREKSLIVATTSGSEIKNDEITHDKEARK 1020
DB |||||
QY 966 KLKGIPOLEGQDMITWILQMFDTVKRLREKSLIVATTSGSEIKNDEITHDKEARK 1025
DB |||||
QY 1021 RKAARLHRQKIMQMSALQNFITETKLMYDNTSEMPGKEDSIWEEESTPAVDSYRI 1080
DB |||||
QY 1026 RKAARLHRQKIMQMSALQNFITETKLMYDNTSEMPGKEDSIWEEESTPAVDSYRI 1085
DB |||||
QY 1081 ALGPKRGSPVTEKVLTCILCOBEQEVKIENNAMVLSACVQKSTALTQHRGKPIELSGEA 1140
DB |||||
QY 1086 ALGPKRGSPVTEKVLTCILCOBEQEVKIENNAMVLSACVQKSTALTQHRGKPIELSGEA 1145
DB |||||
QY 1141 LDPLFMDPLAYGTTGSGHVMHACWQKYPEAVQLSSQQRHVDLPLESGEYLCPIC 1200
DB |||||
QY 1146 LDPLFMDPLAYGTTGSGHVMHACWQKYPEAVQLSSQQRHVDLPLESGEYLCPIC 1205
DB |||||
QY 1201 KSLCNTVPIIPLOPKINSENADALAOILLTARWITQTLARISGYNIRHAKGENPIPI 1260
DB |||||
QY 1206 KSLCNTVPIIPLOPKINSENADALAOILLTARWITQTLARISGYNIRHAKGENPIPI 1265
DB |||||
QY 1261 FNOGMDSTLEPHSILSFVRESSIKYSIKEMVILFATTIYRIGLVKPPDRPRVPM 1320
DB |||||
QY 1266 FNOGMDSTLEPHSILSFVRESSIKYSIKEMVILFATTIYRIGLVKPPDRPRVPM 1325
DB |||||
QY 1321 TWSTCAFTIQAENILGDEKPLFGALQNRHNGKALMQFAVQRIITCPQVLIQHLVR 1380
DB |||||
QY 1326 TWSTCAFTIQAENILGDEKPLFGALQNRHNGKALMQFAVQRIITCPQVLIQHLVR 1385
DB |||||
QY 1381 LLSVLPNPKSBDTPCLLSIDLPFLVGVAVLAPPSLYWDDPVDLPSSSSYNHLYLFH 1440
DB |||||
QY 1386 LLSVLPNPKSBDTPCLLSIDLPFLVGVAVLAPPSLYWDDPVDLPSSSSYNHLYLFH 1445
DB |||||
QY 1441 LITMAHMLQILLTVDGTLPLAQVQEDSEAHSAFFABISQYTSIGCDIPGWYLV 1500
DB |||||
QY 1446 LITMAHMLQILLTVDGTLPLAQVQEDSEAHSAFFABISQYTSIGCDIPGWYLV 1505
DB |||||
QY 1501 LKNGITPYLRCAALPFLYLLGVTTPPEELHTNSAEGEYSALCSYLSLPTNLLFOBYMDT 1560
DB |||||
QY 1506 LKNGITPYLRCAALPFLYLLGVTTPPEELHTNSAEGEYSALCSYLSLPTNLLFOBYMDT 1565
DB |||||
QY 1561 VRPLLQRCADPALLNCLKQNTVVRYPKRNLSLIELPDDYSCLANQASHFRCPSADDE 1620
DB |||||
QY 1566 VRPLLQRCADPALLNCLKQNTVVRYPKRNLSLIELPDDYSCLANQASHFRCPSADDE 1625
DB |||||
QY 1621 RXHPVLCPLCGAILCSQNTCCQEI VNGEYVACIFHALHCGAGVCIPLKIRECRVVLVEG 1680
DB |||||
QY 1626 RXHPVLCPLCGAILCSQNTCCQEI VNGEYVACIFHALHCGAGVCIPLKIRECRVVLVEG 1685
DB |||||
QY 1681 KARGCAYPAPYLDYGETDPGLKRGNPLHLSRERYKHLVWQHCIIIEIARSQETNOM 1740
DB |||||
QY 1666 KARGCAYPAPYLDYGETDPGLKRGNPLHLSRERYKHLVWQHCIIIEIARSQETNOM 1725
DB |||||
QY 1741 LFGFNWQLL 1749
DB |||||
QY 1726 LFGFNWQLL 1734
DB |||||

RESULT 6
ID AAW84351
XX AAW84351 standard; protein; 1757 AA.
AC AAW84351;
XX AAW84351;
XX 25-MAR-1999 (first entry)
DT Murine ubiquitin-protein ligase Ub1.
DE Ubiquitin-protein ligase; Ub1; mouse; ubiquitinylation; degradation;
XX N-end rule pathway; stress-related muscle wasting; inhibitor; screen.
KW Mus sp.
OS US5861312-A.
XX PN

XX 19-JAN-1999.
PD 02-DEC-1997; 97US-00982956.
XX 02-DEC-1997; 97US-00982956.
PR 02-DEC-1997; 97US-00982956.
XX (CALY) CALIFORNIA INST OF TECHNOLOGY.
XX Kwon YT, Varshavsky A;
PI WPI; 1999-130395/11.
DR N-PSDB; AAV99308.
XX Mouse and human Ub1 cDNA - useful for producing recombinant Ub1
PT polypeptides.
XX Disclosure; Col 15-28; 18pp; English.
CC The present sequence represents a ubiquitin-protein ligase called Ub1.
CC The Ub1 enzymes are involved in protein ubiquitinylation and ultimate
CC degradation through the N-end rule pathway and have been linked to stress
CC -related muscle wasting. Recombinant Ub1 polypeptides can be used to
CC screen for inhibitors of muscle wasting when this is associated with the
CC N-end rule pathway
XX N-end rule pathway
XX Sequence 1757 AA;
QY Query Match 93.4%; Score 8617; DB 2; Length 1757;
DB Best Local Similarity 92.1%; Pred. NO. 0;
Matches 1619; Conservative 64; Mismatches 66; Indels 8; Gaps 3;
QY 1 MADEAGATGMEISAEPLQTPORLASWMDQVDFYTAFLHQAQLVPEIYFAEMDDPLE 60
DB 1 MADEEMDGAERMDVSPPEPLAPQRPASWMDQVDFYTAFLHQAQLVPEIYFAEMDDPLE 60
QY 61 KOESVQMSIITPLEWYLFGBDDPDIKLEKXHSAGFOLCGRVKSGETTYSCRDCAIDPT 120
DB 61 KOESVQMSIITPLEWYLFGBDDPDIKLEKXHSAGFOLCGRVKSGETTYSCRDCAIDPT 120
QY 121 CVLQMDCFQDSVHKHRYKMTSTGGFCDCDGTAEWTKGFCVNVNHPGRAGTIKENSRC 180
DB 121 CVLQMDCFQDSVHKHRYKMTSTGGFCDCDGTAEWTKGFCVNVNHPGRAGTIKENSRC 180
QY 181 PLNEEVIQARRIIPPSVIKYIVEMTIWEEKELPELOIREKNERYCYVLFNDEHSHYDH 240
DB 181 PLNEEVIQARRIIPPSVIKYIVEMTIWEEKELPELOIREKNERYCYVLFNDEHSHYDH 240
QY 241 VYLSQALDCELAQAOLHTTAIDKEGRRAKAGAYAAACQAKEDIKSHSENVSOHPLHV 300
DB 241 VYLSQALDCELAQAOLHTTAIDKEGRRAKAGAYAAACQAKEDIKSHSENVSOHPLHV 300
QY 301 EYLHSEIMAHQKAFALRLGSMWNKIMSYSDFRQIFCQACLREPDSENPCLISRLMLWA 360
DB 301 EYLHSEIMAHQKAFALRLGSMWNKIMSYSDFRQIFCQACLREPDSENPCLISRLMLWA 360
QY 361 KLYGARKILHELIFSSPFMEYKYLKLFAMEVVKYKQLOKEYISIDDDHRSISITALSVO 420
DB 361 KLYGARKILHELIFSSPFMEYKYLKLFAMEVVKYKQLOKEYISIDDDHRSISITALSVO 420
QY 421 MPTVPTLARHLIEEQNVISVITETLEVLPEYLDNRNKNFNFGYQSDKLGRYVAVICDLK 480
DB 421 MPTVPTLARHLIEEQNVISVITETLEVLPEYLDNRNKNFNFGYQSDKLGRYVAVICDLK 480
QY 481 YILISKPTIWTLESLMQFLGFRSFLKILTQCMQMEIRROVQGHIEVDPDWEAAIAIQM 540
DB 481 YILISKPTIWTLESLMQFLGFRSFLKILTQCMQMEIRROVQGHIEVDPDWEAAIAIQM 540
QY 541 QLNKILLMFOEWCAEDDELLVAYKECHKAVMRCSTSTSSSKTVVQSGHSLKTSYKV 600
DB 541 QLNKILLMFOEWCAEDDELLVAYKECHKAVMRCSTSTSSSKTVVQSGHSLKTSYKV 600
QY 601 SEDLVSIIHLPLSRTLAGLHVRLGAVSRLEHFVSFEDFQVEVLVEYPLRCLVLVAQV 660
DB |||||

601 SEDLSIHPLSRTLAGLHVRSLRGAISRLHEFPDFSFQVEVLVEYFLRCLVLVAQVV 660
661 AEMWRNGLSLISQVYQDVKREEMDKDIIIMQIGASLMDPNKFLLLVLQRYELAEA 720
661 AEMWRNGLSLISQVYQDVKREEMDKDIIIMQIGASLMDPNKFLLLVLQRYELTDA 720
721 FNKTISTKQDOLIKQNTLIEBMLQVLIYVGERVYVGVNVTKEBVTWREIILLICIEP 780
721 FNKTISTKQDOLIKQNTLIEBMLQVLIYVGERVYVGVNVTKEBVTWREIILLICIEP 780
781 MPHSAIAKLVLPENNETGLENVINKVATPKPGVSGHGVYELKDBSLKDPNNYFHYSK 840
781 MPHSAIAKLVLPENNETGLENVINKVATPKPGVSGHGVYELKDBSLKDPNNYFHYSK 840
841 TQSKAEMHOKRRKQENKDEALPPPPPEFCFAFSKVINLLNCDIMMYILRTVFERAID 900
841 TQSKAEMHOKRRKQENKDEALPPPPPEFCFAFSKVINLLNCDIMMYILRTVFERAID 900
901 TDSNLWTEGMLQWAFHILALGLEEKQOLQKAPBEVTPDFTHKASRLGSSAM---NIOM 957
901 TDSNLWTEGMLQWAFHILALGLEEKQOLQKAPBEVTPDFTHKASRLGSSAMNAQION 960
958 LLEKLGIPQSGQDMITWILQMDTVKRLREKSLIVATTSGSEIKNDSEITHDKKA 1017
961 LLERLKGIPQSGQDMITWILQMDTVKRLREKSLIVATTSGSEIKNDSEITHDKKA 1020
1018 ERKKAERARLHROKIMQMSALQKNFIETHKMTDNTSEMPGKEDSIMEESTPAVS DY 1077
1021 ERKKAERARLHROKIMQMSALQKNFIETHKMTDNTSEMPGKEDSIMEESTPAVS SEA 1080
1078 SERIALGPKRGPVTEKEVLTCLCQEBQEVKINNAMVLSACVQKSTALTQHRGPBILS 1137
1081 SERIALGPKRGPVTEKEVLTCLCQEBQEVKINNAMVLSACVQKSTALTQHRGPBVDHL 1140
1138 GEALDPLMDPLAGTGTGSGHVMHVCWKYFEAVQLSSQQRHVDLDFLEGEVYL 1197
1141 GETLOPLMDPLAGTGTGSGHVMHVCWKYFEAVQLSSQQRHVDLDFLEGEVYL 1200
1198 PLCKSLCNTVPIPILOPKINSENADALAQLLTLARWITQTLARISGYNIRHAKGENP- 1256
1201 PLCKSLCNTVPIPILOPKINSENADALAQLLTLARWITQTLARISGYNIRHAKGEAPA 1260
1257 IPFPNQMGDSTLFHFSILSGVSSIKYSIKEMVILFATTIYRIGLVKVPDPR 1316
1261 VPVLFNQMGDSTLFHFSILSGVSSIKYSIKEMVILFATTIYRIGLVKVPDPR 1320
1317 VPMLTWTSTCAFTIOAIENLLGEGKPLFGALQNRQHNGIKALMQFAVAQRITCPQVLIQK 1376
1321 VPMLTWTSTCAFTIOAIENLLGEGKPLFGALQNRQHNGIKALMQFAVAQRATCPQVLIHK 1380
1377 HLVRLLSVLVPNIKSDPTCLLSIDLPHVLGAVLAPSLYWDPDVLPQSSVSSYNHL 1436
1381 HLVRLLSVLVPNIKSDPTCLLSIDLPHVLGAVLAPSLYWDPDVLPQSSVSSYNHL 1440
1437 YLFHLITWMLQILLTVDY----GLPLAQVEDSEEAHSFPFAEISQYTSIGCDI 1492
1441 YLFHLITWMLQILLTVDY----GLPLAQVEDSEEAHSFPFAEISQYTSIGCDI 1500
1493 PGWYLVWSLXNGITPYLRCAALFFHYLGVTPPPELHTNSARGEVSA LCSYLSTPNLFL 1552
1501 PGWYLVWSLXNGITPYLRCAALFFHYLGVTPPPELHTNSARGEVSA LCSYLSTPNLFL 1560
1553 LPQEWYDVTVRPLQWCAADPALLNCLKOKNTVVRPRKNSLIELPDDYSCLLNQAHPR 1612
1561 LPQEWYDVTVRPLQWCAADPALLNCLKOKNTVVRPRKNSLIELPDDYSCLLNQAHPR 1620
1613 CPRSADDERKHPVLCIFCGAILCSQNICQEI VNGEEVGCACIFHALHCGAGVCIPLKIRE 1672
1621 CPRSADDERKHPVLCIFCGAILCSQNICQEI VNGEEVGCACIFHALHCGAGVCIPLKIRE 1680
1673 CRVILVEGKARGCAYPAPYLDYBGTDPGLKRGNPLHLSRERYRKLHLVWQOHCIIIBETA 1732

Db 1681 CRVILVEGKARGCAYPAPYLDYBGTDPGLKRGNPLHLSRERYRKLHLVWQOHCIIIBETA 1740
QY 1733 RSQETNQMLFGFNWOLL 1749
Db 1741 RSQETNQMLFGFNWOLL 1757
RESULT 7
ID AAB31162 standard; protein; 1757 AA.
XX AAB31162;
AC AAB31162;
XX 02-APR-2001 (first entry)
XX Amino acid sequence of Mouse Ubr1 protein.
XX Ubr1; E3-type protein; ubiquitin system; ubiquitin-protein ligase;
KW N-end rule pathway; intracellular pathogen; Lysteria monocytogenes;
KW Yersinia enterocolitica; muscle wasting; infection.
XX Mus sp.
XX US6159732-A.
XX 12-DEC-2000.
XX 11-JAN-1999; 99US-00228317.
XX 02-DEC-1997; 97US-00982956.
XX (CALY) CALIFORNIA INST OF TECHNOLOGY.
XX Kwon YT, Varshavsky A;
DR WPI. 2001-090278/10.
DR N-PSDB; AAC86933.
XX Inhibiting the N-end rule pathway in mammalian cells for treating
PT infections and various diseases associated with muscle tissue wasting, by
PT inhibiting the expression of Ubr1 gene.
XX Example; Col 15-28; 18pp; English.
XX The present sequence represents a murine Ubr1 enzyme. Ubr1 is an E3-type
CC protein of the ubiquitin system. Specifically, it is a ubiquitin-protein
CC ligase. The enzyme is specific for destabilising residues exposed at the
CC N-terminus of protein substrates. Inhibition of the expression of Ubr1
CC gene in a cell results in inhibition of the N-end rule pathway. The
CC method is used for treatment of mammalian cells infected with an
CC intracellular pathogen, e.g. Lysteria monocytogenes or Yersinia
CC enterocolitica. Inhibition of N-end rule pathway is also useful for
CC treating various diseases associated with wasting of muscle tissue and
CC infections
XX Sequence 1757 AA;
SQ Query Match 93.4%; Score 8617; DB 4; Length 1757;
Best Local Similarity 92.1%; Pred. No. 0;
Matches 1619; Conservative 64; Mismatches 66; Indels 8; Gaps 3;
QY 1 MADEAGATERMEISAEIPQTPQRLASWDDQVDFYTAFLHHLAQLVPEIYFAEMDDPLE 60
Db 1 MADEEMDGAERMDVSPPEPLAPQRPASWDDQVDFYTAFLHHLAQLVPEIYFAEMDDPLE 60
QY 61 KOESVQMSIITPLEWYLFGEEDDPCLEKXHSAPQLCGRVFSGGTTTSCRDCAIDPT 120
Db 61 KOESVQMSIITPLEWYLFGEEDDPCLEKXHSAPQLCGRVFSGGTTTSCRDCAIDPT 120
QY 121 CVLCHDCQDSVHNHRYKMTSTGGGFCDCGDEATKTPGVNHEPGRAGTIKENSRC 180
Db 121 CVLCHDCQDSVHNHRYKMTSTGGGFCDCGDEATKTPGVNHEPGRAGTIKENSRC 180
121 CVLCHDCQSSVHNHRYKMTSTGGGFCDCGDEATKTPGVNHEPGRAGTIKENSRC 180

181 PLNEEVIVQARKIPSPVIVVEMTIBEEKEPPELOIREKNERYCYVFNDEHSHYD 240
181 PLNEEVIAQARRIFPSVIVVEMTIBEEKEPPELOIREKNERYCYVFNDEHSHYD 240
241 VIYS-ORALDCLABLAQHHTTAIDKEGRRAVAGAYAAQCEAKEDIKSHSENVSHPLHV 300
241 VIYS-ORALDCLABLAQHHTTAIDKEGRRAVAGAYATCQEAKEEDIKSHSENVSHPLHV 300
301 EYLHSEIWAHQFALRIGSMWNKINSYSDPFOIFCOACLEPSENPCLISRLMLWDA 360
301 EYLHSEIWAHQFALRIGSMWNKINSYSDPFOIFCOACLEPSENPCLISRLMLWDA 360
361 KUYKGARKILHLEIIPSSPMEMEKYKLPAMEFVKYKLOKEIYISDDHRSISITALSVO 420
361 KUYKGARKILHLEIIPSSPMEMEKYKLPAMEFVKYKLOKEIYISDDHRSISITALSVO 420
421 MFTVPTLARHLIEEQNVISVITETILLEVPEYLDNRNKFNFQYSGDKLGRVYVICDLK 480
421 MFTVPTLARHLIEEQNVISVITETILLEVPEYLDNRNKFNFQYSGDKLGRVYVICDLK 480
481 YLISKPPTWTERLRMOFLBGRSPFLKILTCMOGHEIRROVGOHIEVDPDWEAAIAIOM 540
481 YLISKPPTWTERLRMOFLBGRSPFLKILTCMOGHEIRROVGOHIEVDPDWEAAIAIOM 540
541 QLNKILLMFOEWCACDEBELLVAYKECHKAVNRCTSPFISSKTVVQSGHSLKTSYRV 600
541 QLNKILLMFOEWCACDEBELLVAYKECHKAVNRCTSPFISSKTVVQSGHSLKTSYRV 600
601 SEDLVSIHILPLSRTLAGHLVRLSRLGAVSRLEHFEVSPFQVEVLVEYPLRCLVLVAQV 660
601 SEDLVSIHILPLSRTLAGHLVRLSRLGAVSRLEHFEVSPFQVEVLVEYPLRCLVLVAQV 660
661 AEMWRNGLSLISQVFTYQDVKCREMYDKOIMLOIGASLMDPKFLLVLQRYELABA 720
661 AEMWRNGLSLISQVFTYQDVKCREMYDKOIMLOIGASLMDPKFLLVLQRYELABA 720
721 FNKTIISTKODDILIKOYNTLIEBMLQVLIYIGERVPGVGNVTKBEVTMREIHLICIEP 780
721 FNKTIISTKODDILIKOYNTLIEBMLQVLIYIGERVPGVGNVTKBEVTMREIHLICIEP 780
781 MPHSIAKALPENENNETGLENVINKVATFKKPGVSGHGVYELKDESLKDFNMFYHYSK 840
781 MPHSIAKALPENENNETGLENVINKVATFKKPGVSGHGVYELKDESLKDFNMFYHYSK 840
841 TQHSKAEHMQKRRQKQENKDEALPPPPPPPCFPAFSKVNNLSCDVMYILTIPIERAVD 900
841 TQHSKAEHMQKRRQKQENKDEALPPPPPPPCFPAFSKVNNLSCDVMYILTIPIERAVD 900
901 TDSNLWTEGMLQMAFHILALGLEKQLOKAPPEBEVTFDFYHKASRLGSSAM--NIQM 957
901 TDSNLWTEGMLQMAFHILALGLEKQLOKAPPEBEVTFDFYHKASRLGSSAM--NIQM 957
958 LLEKLGIPQLEGQKDMITWILQMFDTVKRLREKSLIVATTSGSESIKNDIEITHDKKA 1017
958 LLEKLGIPQLEGQKDMITWILQMFDTVKRLREKSLIVATTSGSESIKNDIEITHDKKA 1017
1018 ERKRAEAAARLHRQKIMQMSALQKNPIETHKMLDNTSEMPGKEDSIMEBESTPAVS DY 1077
1018 ERKRAEAAARLHRQKIMQMSALQKNPIETHKMLDNTSEMPGKEDSIMEBESTPAVS DY 1077
1021 ERKRAEAAARLHRQKIMQMSALQKNPIETHKMLDNTSEMTGKEDSIMEBESTSAVSEA 1080
1021 ERKRAEAAARLHRQKIMQMSALQKNPIETHKMLDNTSEMTGKEDSIMEBESTSAVSEA 1080
1078 SERIALGPKGSPVTEKEVLTICLCOBEQVKIENNAMVLSACVOKSTALTQHRGKPIELS 1137
1078 SERIALGPKGSPVTEKEVLTICLCOBEQVKIENNAMVLSACVOKSTALTQHRGKPIELS 1137
1081 SERIALGPKGPAVTEKEVLTICLCOBEQVKIENNAMVLSACVOKSTALTQHRGKPDHIL 1140
1081 SERIALGPKGPAVTEKEVLTICLCOBEQVKIENNAMVLSACVOKSTALTQHRGKPDHIL 1140
1138 GERALDPLFMDPLAYTYTSGCHVNHAVCWQYFEAVQLSSQQRHIVDLFDLESGEYLC 1197
1138 GERALDPLFMDPLAYTYTSGCHVNHAVCWQYFEAVQLSSQQRHIVDLFDLESGEYLC 1197
1141 GETLDPLFMDPLAHGTYTSGCHVNHAVCWQYFEAVQLSSQQRHIVDLFDLESGEYLC 1200
1141 GETLDPLFMDPLAHGTYTSGCHVNHAVCWQYFEAVQLSSQQRHIVDLFDLESGEYLC 1200
1198 PLCKSLCNTVPIPIPIPOKINSNADALAQLLTLARWITQTLARISGNINIRHAKGNP- 1256
1198 PLCKSLCNTVPIPIPIPOKINSNADALAQLLTLARWITQTLARISGNINIRHAKGNP- 1256
1201 PLCKSLCNTVPIPIPIPOKINSNADALAQLLTLARWITQTLARISGNINIRHAKGEAPA 1260
1201 PLCKSLCNTVPIPIPIPOKINSNADALAQLLTLARWITQTLARISGNINIRHAKGEAPA 1260
1257 IPFFNQMGMDSTLEPHSILSFGVSSIKVNSIKEMVILPATTIYRIGLKVPPDERDPR 1316
1257 IPFFNQMGMDSTLEPHSILSFGVSSIKVNSIKEMVILPATTIYRIGLKVPPDERDPR 1316

1261 VPVLFNQMGMDSTLEPHSILSFGVSSIKVNSIKEMVILPATTIYRIGLKVPPDERDPR 1320
1317 VPMLTWTSTCAPTIQAIENLLGDECKPLFGALQNRHNGLKALMQFAVQARTTCPOVLIQK 1376
1321 VPMTWTSTCAPTIQAIENLLGDECKPLFGALQNRHNGLKALMQFAVQARTTCPOVLIHK 1380
1377 HLVRLLSVLFPNPKSDETPCLLSIDLPHVLVGVAVLAPPSLYWDDPVDLQPSVSSVSNHL 1436
1381 HLARLLSVLFPNPKSDETPCLLSIDLPHVLVGVAVLAPPSLYWDDPVDLQPSVSSVSNHL 1440
1437 YLPHLITMAHMLQILLTVDT---GLPLAQVQEDSERAHSSAPFAEISQVTSIGCDI 1492
1441 YLPHLITMAHMLQILLTVDT---GLPLAQVQEDSERAHSSAPFAEISQVTSIGCDI 1500
1493 PGWTLVSLKNGITPYLRCAALPHYLILGVTPPELHTNSAGEYSALCSVLSLPTNLFL 1552
1501 PGWTLVSLKNGITPYLRCAALPHYLILGVTPPELHTNSAGEYSALCSVLSLPTNLFL 1560
1553 LFOYWDVTVRPLQRCWCPADPALLNCLQKNTVVRPKRNSLIELPDDYSCLLMQASHFR 1612
1561 LFOYWDVTVRPLQRCWCPADPALLNCLQKNTVVRPKRNSLIELPDDYSCLLMQASHFR 1620
1613 CPRSADDERKHPVLCFLCGAILCSQNTCCQBIVNGEVEGACIFHALHCGAGVCIFLKIRE 1672
1621 CPRSADDERKHPVLCFLCGAILCSQNTCCQBIVNGEVEGACIFHALHCGAGVCIFLKIRE 1680
1673 CRVVLVEGKARGCAYPAPYLDYGETDPLKRGKGNPLHLSRERYKHLVWQHCIIIEBIA 1732
1681 CRVVLVEGKARGCAYPAPYLDYGETDPLKRGKGNPLHLSRERYKHLVWQHCIIIEBIA 1740
1733 RSQETNQMLPGFNWOLL 1749
1741 RSQETNQMLPGFNWOLL 1757
RESULT 8
ADJ95455
ID ADJ95455 standard; protein; 1757 AA.
XX
AC ADJ95455;
XX
DT 03-JUN-2004 (first entry)
XX Mouse Ubiquitin ligase E3alpha I.
XX
XX Mouse; enzyme; ubiquitin ligase; E3alpha I; ubiquitin-proteasome pathway;
KW gene therapy; vaccine; muscular atrophy; cachexia; catabolic disorders;
KW cancer cachexia; renal cachexia; inflammatory cachexia;
KW muscle wasting disorder; metabolic acidosis; uremia; burn;
KW hyperthyroidism; Cushing's syndrome; fasting; denervation atrophy;
KW diabetes mellitus; sepsis; AIDS wasting syndrome.
XX Mus musculus.
OS
XX US6706505-B1.
XX
XX 16-MAR-2004.
XX
XX 28-NOV-2000; 2000US-00724126.
XX
XX 08-MAR-2000; 2000US-0187911P.
XX
XX (AMGE-) AMGEN INC.
XX
XX Han H, Kwak K;
XX
XX WPI; 2004-236723/22.
XX
XX New nucleic acid molecule, useful for preparing a composition for
PT diagnosing, treating or preventing diseases associated with human
PT E3approximately polypeptide, e.g., muscle atrophy.
XX

PS Example 1; SEQ ID NO 15; 104pp; English.

XX The invention relates to a new isolated nucleic acid molecule appearing
CC as ADJ95441 (or its complement) encoding a ubiquitin ligase E3alpha I
CC protein appearing as ADJ95442. Also included are a vector comprising the
CC nucleic acid, a host cell comprising the vector, a process of producing a
CC E3alpha I ubiquitin ligase polypeptide, a composition comprising the
CC nucleic acid molecule, a reagent comprising a detectably labelled
CC nucleotide, and a method for determining the presence of a human E3alpha
CC I ubiquitin ligase nucleic acid in a biological sample. The nucleic acid
CC molecule is useful for preparing a composition for diagnosing, treating
CC or preventing diseases associated with human E3alpha I polypeptide, e.g.
CC muscle atrophy, cachexia, catabolic disorders, cancer cachexia, renal
CC metabolic acidosis, uremia, burns, hyperthyroidism, Cushing's syndrome,
CC fasting, denervation atrophy, diabetes mellitus, sepsis and AIDS wasting
CC syndrome. The present sequence represents mouse E3alpha I.

XX Sequence 1757 AA;

Query Match 93.4%; Score 8617; DB 8; Length 1757;

Best Local Similarity 92.1%; Pred. No. 0;

Matches 1619; Conservative 64; Mismatches 66; Indels 8; Gaps 3;

QY 1 MADEAGGTERMEISNELPQTORLASWDDQVDFYTAFLHHLAQIVPEIYPAENDPDLE 60
DB 1 MADEEMDGAERMDVSPPEPLAPORPASWDDQVDFYTAFLHHLAQIVPEIYPAENDPDLE 60
QY 61 KOEESVQMSIPTLEWYLFGEHPDCLCKLKHSGAPQLCGRVFKSGGETTYSRCDCAIDPT 120
DB 61 KOEESVQMSILTPLEWYLFGEHPDCLCKLKHSGAPQLCGRVFKSGGETTYSRCDCAIDPT 120
QY 121 CVLQMDPCFQSSVHKHRYKQHTSTGGFCDCGDETAWKTPPCVNHPEGRAGTICKENSRK 180
DB 121 CVLQMDPCFQSSVHKHRYKQHTSTGGFCDCGDETAWKTPPCVNHPEGRAGTICKENSRK 180
QY 181 PLNEEVIQARIPKPSVIKYVEMTIWEKEKLPPELOIREKNERYCYVLFNDEHSHYDH 240
DB 181 PLNEEVIQARIPKPSVIKYVEMTIWEKEKLPPELOIREKNERYCYVLFNDEHSHYDH 240
QY 241 VIYSLQALDCLAEALHTTAIDKEGRAVAKAGAYACORAKEDIKSHSENVSOHPLHV 300
DB 241 VIYSLQALDCLAEALHTTAIDKEGRAVAKAGAYATCOEAKEDIKSHSENVSOHPLHV 300
QY 301 EYLHSEIMAHQKFAFLRLGSMWNKIMSYSSDFRQIFCOACLEBEPDSENPCLISRLMLMDA 360
DB 301 EYLHSEIMAHQKFAFLRLGSMWNKIMSYSSDFRQIFCOACLEBEPDSENPCLISRLMLMDA 360
QY 361 KLYGARKILHLIELFSSPFMEYKCLFAMEFVKYKQLQKEYISDDHRSISITALSQV 420
DB 361 KLYGARKILHLIELFSSPFMEYKCLFAMEFVKYKQLQKEYISDDHRSISITALSQV 420
QY 421 MFTVPTLARHLIERQNVISVITETLEVLPEYLDNRNKNFPOGYSODKLGRIYAVICDLK 480
DB 421 MFTVPTLARHLIERQNVISVITETLEVLPEYLDNRNKNFPOGYSODKLGRIYAVICDLK 480
QY 481 YILISKPTITWRLRMQFLGFRSFLKILTCMQGHEIRROVQGHIEVDPPWEAAIAIQM 540
DB 481 YILISKPTITWRLRMQFLGFRSFLKILTCMQGHEIRROVQGHIEVDPPWEAAIAIQM 540
QY 541 QLNKILLMPQEWACDEBELLVAYKECHAVNRCSTSFSSKTVVQSGHSLKTSYRV 600
DB 541 QLNKILLMPQEWACDEBELLVAYKECHAVNRCSTSFSSKTVVQSGHSLKTSYRV 600
QY 601 SEDLVSIHLPLSRTAGLHVLSRLGAVSRLEHFEVPEQVLEVEPLRCLVLVAQV 660
DB 601 SEDLVSIHLPLSRTAGLHVLSRLGAVSRLEHFEVPEQVLEVEPLRCLVLVAQV 660
QY 661 AEMWRNGLSLISQVFFYQDVKCREMYDKOIMLQIGASLMDPNKFLLLVLQRYELAEA 720
DB 661 AEMWRNGLSLISQVFFYQDVKCREMYDKOIMLQIGASLMDPNKFLLLVLQRYELAEA 720
QY 721 FNKTISTKDDQLIKQNTLIEBMLQVLIYVGERYVPGVGNVTKEBVTREIHLCLIEP 780

RESULT 9

DB 721 FNKTISTKDDQLIKQNTLIEBMLQVLIYVGERYVPGVGNVTREIHLCLIEP 780
QY 781 MPHSAIAKLNLENENNETGLENNINKVATPKKPGVSGHGVYELKDESLKDNFMYFYHYSK 840
DB 781 MPHSAIAKLNLENENNETGLENNINKVATPKKPGVSGHGVYELKDESLKDNFMYFYHYSK 840
QY 841 TOHSAEHMQKRRKQENKDBALPPPPPPPCFAPSKVINLNLNCIMMILATVFERAID 900
DB 841 TOHSAEHMQKRRKQENKDBALPPPPPPPCFAPSKVINLNLNCIMMILATVFERAID 900
QY 901 TDSNLTWTEGMLQMAFHIALGALLREKQOLQKAPREEVTFDFVHKASRLGSSAM---NLCM 957
DB 901 TDSNLTWTEGMLQMAFHIALGALLREKQOLQKAPREEVTFDFVHKASRLGSSAMNLCM 957
QY 958 LLEKLGIPQLEGQKDMITWILQMPDVTVKRLREKSLIVATTTSGSESINKNDBIETHKEKA 1017
DB 958 LLEKLGIPQLEGQKDMITWILQMPDVTVKRLREKSLIVATTTSGSESINKNDBIETHKEKA 1017
QY 1018 SERIALGKRGPSVTEKEVLTCLCOEEOEVKIENNAMVLSACVOKSTALTQHRGKPIELS 1137
DB 1018 SERIALGKRGPSVTEKEVLTCLCOEEOEVKIENNAMVLSACVOKSTALTQHRGKPIELS 1137
QY 1138 GEALDPLFMDPLDLAGTYTSGCHVMHACVQKYPFAVQLSSQQRHIVDLFDLESSEYLC 1197
DB 1138 GEALDPLFMDPLDLAGTYTSGCHVMHACVQKYPFAVQLSSQQRHIVDLFDLESSEYLC 1197
QY 1201 PLCKSLCNTVPIIPILOPKINSNADALAOULLTARMIQTVLARIQSYNIRHAKGNP- 1256
DB 1201 PLCKSLCNTVPIIPILOPKINSNADALAOULLTARMIQTVLARIQSYNIRHAKGNP- 1256
QY 1257 IPIFNQMGSGSTLEFHSILSPGVSSIKEMVILFATTYRIGLVKPPDERDPR 1316
DB 1257 IPIFNQMGSGSTLEFHSILSPGVSSIKEMVILFATTYRIGLVKPPDERDPR 1316
QY 1317 VPMLTWTSCAPTIOAIENLGLDEGKPLFGALQNRHGLKALMQFAVQRIITCPOVLJQK 1376
DB 1317 VPMLTWTSCAPTIOAIENLGLDEGKPLFGALQNRHGLKALMQFAVQRIITCPOVLJQK 1376
QY 1377 HLVRLLSVLPIKSEDPCLLSIDLPHVLVGAVALPPLSYWDDPVDLQPSVSSSYNHL 1436
DB 1377 HLVRLLSVLPIKSEDPCLLSIDLPHVLVGAVALPPLSYWDDPVDLQPSVSSSYNHL 1436
QY 1437 YLPHLITMAHMLQILLITVDT---GLPLAQVQEDSEEAHSAASSPFAEISQVTSIGCDI 1492
DB 1437 YLPHLITMAHMLQILLITVDT---GLPLAQVQEDSEEAHSAASSPFAEISQVTSIGCDI 1492
QY 1493 PGWYLVSLKNGITPYLRCAALPFHYLIGVTPPELHTNSAEGYSALCSYLSLPTNLP 1552
DB 1493 PGWYLVSLKNGITPYLRCAALPFHYLIGVTPPELHTNSAEGYSALCSYLSLPTNLP 1552
QY 1553 LFOEYDVTVRFLQRCWADPALLNCLKQKNTVVPYRKRNSLIELPDYDSCLLNQASHFR 1612
DB 1553 LFOEYDVTVRFLQRCWADPALLNCLKQKNTVVPYRKRNSLIELPDYDSCLLNQASHFR 1612
QY 1613 CPRSADDERKHPVLCFCGAILCSQNTCCQBIIVNGEYGVACIFHALHCGAGVCIPLKIRE 1672
DB 1613 CPRSADDERKHPVLCFCGAILCSQNTCCQBIIVNGEYGVACIFHALHCGAGVCIPLKIRE 1672
QY 1673 CRVVLVEGKARGCAYPAPYLDYGETDPLGKRGNPLHLRERYRKLHLVWQOHCIIIEBIA 1732
DB 1673 CRVVLVEGKARGCAYPAPYLDYGETDPLGKRGNPLHLRERYRKLHLVWQOHCIIIEBIA 1732
QY 1733 RSQETNQMLPGFNQOLL 1749
DB 1749 RSQETNQMLPGFNQOLL 1757

ADS86877

ID ADS86877 standard; protein; 1757 AA.

XX

AC ADS86877;

XX

DT 16-DEC-2004 (first entry)

XX

DE Murine E3alpha ubiquitin ligase, muE3I protein sequence.

XX

KW Mouse; E3alpha ubiquitin ligase; muE3I; ubiquitin-proteasome pathway; rapid muscle wasting; fasting; metabolic acidosis; muscle degeneration; kidney failure; renal cachexia; uremia; diabetes mellitus; sepsis; AIDS wasting syndrome; cancer cachexia; negative nitrogen balance; burn; Cushing's syndrome; inflammatory cachexia; hyperthyroidism; denervation atrophy; protein/tissue wasting; energy-protein malnutrition; muscle atrophy; gene therapy; enzyme.

KW

XX

OS Mus musculus.

XX

XX US2004185037-A1.

XX

XX 23-SEP-2004.

XX

XX 15-JAN-2004; 2004US-00758672.

XX

XX 08-MAR-2000; 2000US-0187911P.

XX

XX 28-NOV-2000; 2000US-00724126.

XX

XX (HANH/) HAN H.

PA

XX (KWAK/) KWAK K.

XX

XX Han H, Kwak K;

XX

XX WPI; 2004-707854/69.

XX

XX Novel isolated human E3alpha ubiquitin ligase nucleic acid molecule

PT

XX useful for treating and/or preventing renal cachexia or inflammatory

XX

XX cachexia.

XX

XX Example 1; SEQ ID NO 15; 115pp; English.

XX

The present invention relates to new orthologue of human E3alpha ubiquitin ligase, huE3alpha1 and huE3alpha11. Most intracellular proteins are degraded through the ubiquitin-proteasome pathway. Proteins are marked for proteasomal degradation by conjugation of ubiquitin to the protein. Conjugation of the ubiquitin molecule involves the activation by E1 enzyme, subsequent transfer to E2 enzyme, which serves as a carrier, and then interacts with a specific E3 ligase family member. E3 ligase binds to proteins targeted for degradation and catalyses the transfer of ubiquitin from the E2 carrier enzyme to the target protein. E3 ligase determines the specificity of the system. The E3alpha family is the main family of intracellular ligases and is involved in the N-end rule pathway of protein degradation. E3alpha enzyme binds directly to the primary destabilising N-terminal amino acid and catalyses ubiquitin conjugation thereby targeting the protein for degradation. The human E3alpha gene is located on chromosome 15 q. Increased proteolysis through the ubiquitin-proteasome pathway has been determined to be a major cause of rapid muscle wasting including, fasting, metabolic acidosis, muscle degeneration, kidney failure, renal cachexia, uremia, diabetes mellitus, sepsis, AIDS wasting syndrome, cancer cachexia, negative nitrogen balance, burns, Cushing's syndrome, inflammatory cachexia, hyperthyroidism, denervation atrophy, protein/tissue wasting, and energy-protein malnutrition. E3alpha plays a role in the overall increase in ubiquitination that is associated with and may mediate muscle atrophy in catabolic and other disease states. Treatment may be administered by gene therapy, cell therapy and antisense therapy methods. The present sequence is murine E3alpha ubiquitin ligase, muE3I protein sequence.

XX

SQ Sequence 1757 AA;

Query Match

Best Local Similarity 93.4%; Score 8617; DB 8; Length 1757;

Matches 1619; Conservative 64; Mismatches 66; Indels 8; Gaps 3;

```
QY 1078 SRALGPKEGSPVTEKEVLTCLCOBEQVKLENNAMVLSACVOKSTALTQHRGPIELS 1137
Db 1081 SRALGPKEGPAVTEKEVLTCLCOBEQVKLENNAMVLSACVOKSTALTQHRGPVDHL 1140
QY 1138 GEALDPLFMDPOLAYGTYTGSCHVNHAVCWQKYFEAVQLSSQORIHVDLFDLESGEYLC 1197
Db 1141 GETLDPLFMDPLANGTYTGSCHVNHAVCWQKYFEAVQLSSQORIHVDLFDLESGEYLC 1200
QY 1198 PLCKSLCNTVPIPILOPKINSNADALAQILLTARWTQTVLARIISGVNIRHAKGNP- 1256
Db 1201 PLCKSLCNTVPIPILOPKINSNADALAQILLTARWTQTVLARIISGVNIRHAKGEADA 1260
QY 1257 IPIFNQMGDSTLEPHSILSCGVSSIKYSIKEMVILFATTIYRIGLKVPDPDRDR 1316
Db 1261 VPIFNQMGDSTFEPHSILSCGVSSIKYSIKEMVILFATTIYRIGLKVPDPDELDR 1320
QY 1317 VPMWTSTCAFTIQAENLLGDEKPLFGALQNRHNGIKALMQFAVQRIICPOVLIQK 1376
Db 1321 VPMWTSTCAFTIQAENLLGDEKPLFGALQNRHNGIKALMQFAVQRIICPOVLIHK 1380
QY 1377 HLVRLSVVLPNIKSEDTPCLLSIDLFHVLGAVLAPPSLYWDDPVDLPQSSVSSYNHL 1436
Db 1381 HVARLLSVLPNLQSENTPGLLSVDLFHVLGAVLAPPSLYWDDTVDLQSPSLSSYNHL 1440
QY 1437 YLPHLITWAHMLQILLTVDT----GLPLAQVEDSEEAHSAPFAEISQYTSIGCDI 1492
Db 1441 YLPHLITWAHMLQILLTVDTDLSGPPLEAGBEDSEARCAAFVVEQSQHTDGLTGCA 1500
QY 1493 PGWYLVSVLKNGITPYLRCAALPFHYLLGVTPPEELHTNSARGESYALCSYLSLPTNLP 1552
Db 1501 PGWYLVSVLKNGITPYLRCAALPFHYLLGVTPPEELFANSARGESYALCSYLSLPTNLP 1560
QY 1553 LFQEWYDVTVRLLQRCWADPALLCNLCQKNTVYRPRKNSIELPDDYSCILNQASHPR 1612
Db 1561 LFQEWYDVTVRLLQRCWADPALLCNLCQKNTVYRPRKNSIELPDDYSCILNQASHPR 1620
QY 1613 CPRSADDERKHPVLCFCAAILCSQNICCOEIVNGEVEGACIFHALHCGAGVCIPLKIRE 1672
Db 1621 CPRSADDERKHPVLCFCAAILCSQNICCOEIVNGEVEGACIFHALHCGAGVCIPLKIRE 1680
QY 1673 CRVVLVEGARGCAVPAPVLDYGETDGLKRGNPILHSRRERYKHLVWQOHCIIERTA 1732
Db 1681 CRVVLVEGARGCAVPAPVLDYGETDGLKRGNPILHSRRERYKHLVWQOHCIIERTA 1740
QY 1733 RSQETNQMLFGFNWQLL 1749
Db 1741 RSQETNQMLFGFNWQLL 1757

RESULT 10
ADI16325
ID ADI16325 standard; protein; 908 AA.
XX
AC ADI16325;
XX
DT 22-APR-2004 (first entry)
XX
DE Human protein modification and maintenance molecule (PMMW) protein #10.
XX
KW human; protein modification and maintenance molecule; PMMW;
KW gastrointestinal disorders; peptic ulcer; Crohn's disease;
KW cardiovascular disorders; hypertension; congenital heart disease;
KW autoimmune disease; inflammatory disease; AIDS; anaemia;
KW developmental disorder; Cushing's syndrome; tubular acidosis;
KW epithelial disorder; eczema; scabies; neurological disorder;
KW Alzheimer's disease; multiple sclerosis; infection; cancer.
XX
OS Homo sapiens.
XX
PN WO2003100016-A2.
XX
XX 04-DEC-2003.
XX
PD
```

```
PF 22-MAY-2003; 2003WO-US016498.
XX
PR 22-MAY-2002; 2002US-0383491P.
PR 24-JUN-2002; 2002US-0391378P.
PR 22-JUL-2002; 2002US-0397921P.
XX
PA (INCY-) INCYTE CORP.
XX
PI Khare R, Bulloch SA, Swarnakar A, Elliott VS, Marquis JP,
PI Mason PM, Chawla NK, Ramkumar J, Kable AE, Hafalia AJA, Lee SY,
PI Tran UK, Yue H, Becha SD, Griffin JA, Chang H, Jiang X, Jackson AA,
PI Richardson TW, Lal PG, Yao MG, Lu Y, Warren BA, Jin P, Wilson AD,
PI Gietzen KJ;
XX
DR WPI; 2004-035124/03.
XX N-PSDB; ADI16377.
XX
DR New protein modification and maintenance molecules, useful for diagnosing
PT or treating e.g. peptic ulcer, hypertension, rheumatic fever, AIDS,
PT Cushing's syndrome, Alzheimer's disease, multiple sclerosis, stroke or
PT cancers.
XX
PS Claim 1; SEQ ID NO 10; 419pp; English.
XX
CC The invention comprises the amino acid and coding sequences of human
CC protein modification and maintenance molecules (PMMW). The DNA and
CC protein sequences of the invention are useful for the diagnosis and
CC treatment of disorders associated with expression of PMMW, such as:
CC gastrointestinal disorders (e.g. peptic ulcer and Crohn's disease),
CC cardiovascular disorders (e.g. hypertension and congenital heart
CC disease), autoimmune or inflammatory disease (e.g. AIDS and anaemia),
CC developmental disorders (e.g. Cushing's syndrome and tubular acidosis),
CC epithelial disorders (e.g. eczema and scabies), neurological disorders
CC (e.g. Alzheimer's disease and multiple sclerosis), infections and cancer.
CC The present amino acid sequence represents a human PMMW protein of the
CC invention.
XX
SQ Sequence 908 AA;
Query Match 51.7%; Score 4768; DB 8; Length 908;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 902; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 12 MEISAEALPQTPQRLASWMDQVDFYAFHLHLAQLVPEIYFAEMDPDLEKQESVQMSIF 71
Db 1 MEISAEALPQTPQRLASWMDQVDFYAFHLHLAQLVPEIYFAEMDPDLEKQESVQMSIF 60
QY 72 TPLEWYLFGEPPDICLEKLKHSAGAFQLCGRVFKSGETTYSRCDCAIDPTCVLCMDCFQDS 131
Db 61 TPLEWYLFGEPPDICLEKLKHSAGAFQLCGRVFKSGETTYSRCDCAIDPTCVLCMDCFQDS 120
QY 132 VHKHRYKMHSTTGGGFCDCGDTBAWTKGPPCVNHEPGRAGTIKENSRCPLNEEVIVQAR 191
Db 121 VHKHRYKMHSTTGGGFCDCGDTBAWTKGPPCVNHEPGRAGTIKENSRCPLNEEVIVQAR 180
QY 192 KIPFSVIKYVEMTITWEEKELPPELQIREKNERYCYLVFNDEHSHSDHYVLSQALDC 251
Db 191 KIPFSVIKYVEMTITWEEKELPPELQIREKNERYCYLVFNDEHSHSDHYVLSQALDC 240
QY 252 ELAEALQHTTAIDKEGRRAVAGAYACQAEKEDIKSHSENVVSHPLHVEVLHSHMAHQ 311
Db 241 ELAEALQHTTAIDKEGRRAVAGAYACQAEKEDIKSHSENVVSHPLHVEVLHSHMAHQ 300
QY 312 KPALRLGSMWNKIMSYSDPRQIFCOACLRREPSENPCLISRLMLWDAKLYKGARKILH 371
Db 301 KPALRLGSMWNKIMSYSDPRQIFCOACLRREPSENPCLISRLMLWDAKLYKGARKILH 360
QY 372 ELIFSSPFMEYKYLKLFAMEFVKYKQIQKEYISDDHDSISITALSVMQFTVPTLARHL 431
Db 361 ELIFSSPFMEYKYLKLFAMEFVKYKQIQKEYISDDHDSISITALSVMQFTVPTLARHL 420
QY 432 IEQNQVSVITETLLEVLPEYLDNRNKNFPGYSGDKLGRVAVTCDLKYLISKPTIWT 491
```

Db 421 IBEQNVISVITETLLEVLPEYLDNRNKNFQGYSDQLGRVAVVAVICDLKYLISKPTIWT 480
 QY 492 ERLRQFLGFSFLKILTCMQGMEIRROVGOHLEVPDWEAAIAIQMLKNILLMFOE 551
 Db 481 ERLRQFLGFSFLKILTCMQGMEIRROVGOHLEVPDWEAAIAIQMLKNILLMFOE 540
 QY 552 WCACDELLVAYKECHKAVMRCSTSFISSTKVVQSCGHSLETYSYRVEDLVSIHLPL 611
 Db 541 WCACDELLVAYKECHKAVMRCSTSFISSTKVVQSCGHSLETYSYRVEDLVSIHLPL 600
 QY 612 SRTLAGLHVLRLGAVSRHLEFVDFEDQVEVLVEYPLRCLVLVAQVVAEMRRNGLSL 671
 Db 601 SRTLAGLHVLRLGAVSRHLEFVDFEDQVEVLVEYPLRCLVLVAQVVAEMRRNGLSL 660
 QY 672 ISQVFFYQDVKCREEMDKDIIIMLQIGASLMDPNKFLILLVQLRYELAEAFNKTISTKDD 731
 Db 661 ISQVFFYQDVKCREEMDKDIIIMLQIGASLMDPNKFLILLVQLRYELAEAFNKTISTKDD 720
 QY 732 LTKQYNTLIEMLQVLIYVGERYVPGVGNVTKKEVTVREIITHLLCIEPMPHSAIAKNLP 791
 Db 721 LTKQYNTLIEMLQVLIYVGERYVPGVGNVTKKEVTVREIITHLLCIEPMPHSAIAKNLP 780
 QY 792 ENNETGLNVINKVATPKPGVGHGYELKDSLKDFNMFYHYSKTQHSKAEHMOK 851
 Db 781 ENNETGLNVINKVATPKPGVGHGYELKDSLKDFNMFYHYSKTQHSKAEHMOK 840
 QY 852 KRRQENKDEALPPPPPPPPFCFAPSKVINLNCDIMMYILRTVPERAIDTDSNLATGKL 911
 Db 841 KRRQENKDEALPPPPPPPPFCFAPSKVINLNCDIMMYILRTVPERAIDTDSNLATGKL 900
 QY 912 QMAFHI 917
 Db 901 QMDWEV 906
 RESULT 11
 ADJ95444
 ID ADJ95444 standard; protein; 1755 AA.
 AC ADJ95444;
 XX
 DT 03-JUN-2004 (first entry)
 DE Human Ubiquitin ligase E3alpha II.
 XX
 KW Human; enzyme; ubiquitin ligase; E3alpha II;
 KW ubiquitin-proteasome pathway; gene therapy; vaccine; muscular atrophy;
 KW cachexia; catabolic disorders; cancer cachexia; renal cachexia;
 KW inflammatory cachexia; muscle wasting disorder; metabolic acidosis;
 KW uremia; burn; hyperthyroidism; Cushing's syndrome; fasting;
 KW denervation atrophy; diabetes mellitus; sepsis; AIDS wasting syndrome.
 XX
 OS Homo sapiens.
 XX
 EN US6706505-B1.
 XX
 PD 16-MAR-2004.
 XX
 PF 28-NOV-2000; 2000US-00724126.
 XX
 PR 08-MAR-2000; 2000US-0187911P.
 XX
 PA (AMGE-) AMGEN INC.
 XX
 PI Han H, Kwak K;
 XX
 DR WPI: 2004-236723/22.
 DR N-PSDB; ADJ95443.
 XX
 XX New nucleic acid molecule, useful for preparing a composition for
 PT diagnosing, treating or preventing diseases associated with human
 PT E3approximate polypeptide, e.g., muscle atrophy.
 XX

PS Example 2; SEQ ID NO 4; 104pp; English.
 XX
 CC The invention relates to a new isolated nucleic acid molecule appearing
 CC as ADJ95441 (or its complement) encoding a ubiquitin ligase E3alpha I
 CC protein appearing as ADJ95442. Also included are a vector comprising a
 CC nucleic acid, a host cell comprising the vector, a process of producing a
 CC E3alpha I ubiquitin ligase polypeptide, a composition comprising the
 CC nucleic acid molecule, a reagent comprising a detectably labelled
 CC nucleotide, and a method for determining the presence of a human E3alpha
 CC I ubiquitin ligase nucleic acid in a biological sample. The nucleic acid
 CC molecule is useful for preparing a composition for diagnosing, treating
 CC or preventing diseases associated with human E3alpha I polypeptide, e.g.
 CC muscle atrophy, cachexia, catabolic disorders, cancer cachexia, renal
 CC cachexia, inflammatory cachexia, muscle wasting disorders associated with
 CC metabolic acidosis, uremia, burns, hyperthyroidism, Cushing's syndrome,
 CC fasting, denervation atrophy, diabetes mellitus, sepsis and AIDS wasting
 CC syndrome. The present sequence represents human E3alpha II.
 XX
 SQ Sequence 1755 AA;
 Query Match 46.9%; Score 4327; DB 8; Length 1755;
 Best Local Similarity 47.2%; Pred. No. 0;
 Matches 841; Conservative 339; Mismatches 543; Indels 58; Gaps 27;
 QY 1 MADE---EAGTER--MEISAEPLQTPORLASWMDQVDFYTAFLHLHLAQVLPRIYPAEM 55
 Db 1 MASELEPEVQAIKSLLECSAE-----EIAAGKWLQATDLTREVVQHLAHVVPKIYCRGP 54
 QY 56 DPDLKQESVQMSIFPTLEWYLFGEHDDPDIKLEKHKSG-AFQLGGRVFKSGEYTSYCRD 114
 Db 55 NPPFKEDMLAQHVLGLGFMWYLCGEDPAFPPKLEQANKPSHLGCRVFKYGEYTSYCRD 114
 QY 115 CAIDPTCVLNDQFQDSVHKHRYKMTSTGGGFCDCGTTEAWTKGPCVNVHPEGRAG-T 173
 Db 115 CAVDPTCVLNCMECFGLSIHRDHYRMTTSGGGGFCDCGTTEAWKEGPGYQCKHELTSTIE 174
 QY 174 IKENSRCPLNEBEVIVQARKIPPSVIKYVVENTIWEKEKLPPELQIIEKNERYCYLVFND 233
 Db 175 EEDPLVHLSDEVARTYNTNIFATPRVAVEILTWEKSELPADELMWEEKSYTYCMLFND 234
 QY 234 EHSYDVHYISLQALDCELAQAQHTTAIDKEGRRAVAGAYACQAKEDDIKSHSNV 293
 Db 235 EVHTYEQVIYTLQKAVNCTQKEAIGFATTVDRGRSRYGDFQYCEQAKSVIRNTSRQ 294
 QY 294 SQHPLHVEVLHSETMAHQKPAIRLGSWMNKIWSYSSDFRQIFCQACLEEDPSNPCLIS 353
 Db 295 TK-PLKQVQMHSSIVAHQNFGLKLLSLWISIGYSDGLRRLILCQVLQEGDGENSSLVD 353
 QY 354 RLMLMDAKLYKGARKILHELIFSSFFMEMEYKCLFAMEFVKYKQLQKEYISDDHDSIS 413
 Db 354 RLMLSDSKLWKGARSVYHQLFWSSILLMDLKYKLPFAVFAKNYQQLQDFMEDDHERAVS 413
 QY 414 ITALSVMQFTVPTLARHLIEQNVISVITETLLEVLPEYLDNRNKNFQGYV---QDKLG 470
 Db 414 VTALSVPQFTAPTARMLITEENLMSIIKTFFMDHL-RHRDAQGRFQERYTALQAPKR 472
 QY 471 RVAVICDLKYLISKPTIETRLRMQFLEGRFLKILTCMQGMEIRROVGOHLEVPD 530
 Db 473 RVQSLILDLYLISKPTESDELRLQFLEGFDAFLLELLKCMQGMDFITRQVGOHIEHEP 532
 QY 531 DWEAAIAIQMLKNILLMFOEWACDEBELLVAYKECHKAVMRCSTSFISSTKVVQS-C 589
 Db 533 EWEAAFTLQMKLTHVISMMQDCASDEKVLIEYKCLAVLMQCHGGYTDGEQPTLSIC 592
 QY 590 GHSLETYSYRVEDLVSIHLPLSRTLAGLHVLRLGAVSRHLEFVDFEDQVEVLVEY 649
 Db 593 GHSVETIRYCVSQEKSVIHLFVSRLLAGLHVLRLGAVSRHLEFVDFEDQVEVLVEY 652
 QY 650 LRLCVLVAQVVAEMRRNGLSLISQVFFYQDVKCREEMDKDIIIMLQIGASLMDPNKFL 709
 Db 653 LRLCVLVAQVVAEMRRNGLSLISQVFFYQDVKCREEMDKDIIIMLQIGASLMDPNKFL 712
 QY 710 LVLQRYELAEAFN-----KTISTK--DQDLIKQYNTLIEEMLQVLIYVGERYVPGVGNV 762

DB 713 ILSRFLQIESTPDYGRFSEITHKDVQVQNNLTIEEMLYLIIMLVGERFSPGVGV 772
QY 763 TKEEVTRMREIIHLLCIEMPHSAIAKNLPENENNETGLENVINKVATFKKPGVSGHGYE 822
DB 773 NATDEIKREIIHQLSIKPMAHSELVKSLPEDENKETGMESVIEVAHFKKPGLTGRGMYE 832
QY 823 LKDESILKDFMVPYHYKTOHSAEKHMOKRKRQENKDEALPPPPPEFCAPSKVINLL 882
DB 833 LKPECAKBNLYFYHFSRAEQKAEBAORLKRQREDTALPPVLPFPFCLFASLVNLT 892
QY 883 NCDIMMYILRTVERAIDTDSNLTGMLQMAFHIALGLLEKQOQKAPREE-VTFDP 941
DB 893 QSDVLMCINGTILQWAVEHNGVANSSEMLQVULHLLGMALEKQKQLENTEHVTF 952
QY 942 YHKASRLGSSAMN---IQMLLEKLGIPQLEGOKMITWILQWFDVTKRLREKS-CLIVA 997
DB 953 TQKISKPGKAPKNSPGLAMLETQNAPLYEVHKOMIRWLKTFNAVKMRESSPSPA 1012
QY 998 TTSSESINKDEITHDKKAEKRAKBAARLHRQKIMQMSALQKNFIETHKLMYDNTSE 1057
DB 1013 ETEGTTIM---BESSRDKAERKRAKBAARLHRQKIMQMSALQKNFIETHKLMYDNTSE 1069
QY 1058 MPKEDSIMBESSTPAVSYSRIALGPKGSPVTEKEVLTCLCOEQEVKLNNAMVLS 1117
DB 1070 LDASTSAVLJDH--SPVASDMLTALGPTQTQVPEQRFVTCILCOEQEVKVESRAMVLA 1127
QY 1118 ACQKSTALTOHRGKPIELSGALDPLFMDPOLAYGTYTSGCHVHVAWCWKYFEAVOL 1177
DB 1128 AFQSTVLKRSKRFQI-DPEKYDPLFMDPOLAYGTYTSGCHVHVAWCWKYFEAVOL 1186
QY 1178 SSOQ-----RIHVDLPDESBEYCLPCKSLQNTVPIIPLOPKINSADALQALLT 1232
DB 1187 KEQRRQORLRLATS-YDVENGEPCLCELSNTVPIPL-LPPNIFNNRLN-FSDOPNL 1243
QY 1233 ARWQIVLARIISGYNIRHAKGNPIPIFNQCGDSTLEFHSILSPGVESSIKYNSIKE 1292
DB 1244 TQWIRTISQIKALQRLKEESTPNNAKNSGVNDELQPLGPRFPDPFKPIPYSESIKE 1303
QY 1293 MVLFATTYIRIGLKVPPDERPRVPMLTWSTCAFTQIAENLLGDGPKPLFGALQNRQH 1352
DB 1304 MLTTFATYKVLGVKHPNEPRVPMCMGSCAYTIQSIERILSDKPLFGPLPCRLD 1363
QY 1353 NGLKALMQFAVQRITCPQVLIQKHLVRLLSVVLNPKISEDTPCLLSIDLPHVLGAVLA 1412
DB 1364 DCLSLTRFAAAHWTVASVYQGHFKPLFASLVNDSSHEELPCILIDIMPHLLVGLVA 1423
QY 1413 PPSLYWDDPVDLOPSSVSSVHLYLFLHITWAMLOILL---TVDTGPLAQVQEDSEE 1469
DB 1424 PPAQOQD-----FSGISLGTGDLHIFHLVTHAHIQILLTSTCEBNGMD--QENPPCEE 1476
QY 1470 AHSASFPAEISQYTSIGSGICDIP-GWLVVSLKNGITPYLRCAALFPHYLLGVTPPEEL 1528
DB 1477 ESAVALYKTLHQYT-GSALKKEIPSGNHLVRSVRAGIMPLKCSALFHYLVGVSPPDI 1535
QY 1529 HTNSAEYSALCSYLSPLTNLFLFQYWDVTRPQLQWQCADPALLNCKOKNTVVRYP 1588
DB 1536 QV-PGTSHEPLCSYLSPLNNLCLFQENSEIMNSLIESMCRNSEVRYLEGEDAIRYP 1594
QY 1589 RKNSLIELDDYSCLLNOASHFRPSRADDKHPVLCLFGAILCSQNICQBIINVE 1648
DB 1595 RESNKLINLPEDYSSILNOASHFRPSRADDKHPVLCLFGAILCSQNICQBIINVE 1654
QY 1649 EYVACIFALHCGAGVCIELKRECVVLVEKAGCAVPAPVLYDEYGETDPLKRGNPL 1708
DB 1655 DVGACTAHTYSCGSGVIFLRVREQVFLRAGKCFYSPYLDYGETDQGLRAGNPL 1714
QY 1709 HLSRERYKHLVWQOHCITIEIARSQETNQLMFGFNWOLL 1749
DB 1715 HLCERFKIKLWQHSVTEIGHAQEANOQLVGLDQHL 1755

RESULT 12

Query Match

46.9%; Score 4327; DB 8; Length 1755;

ADS86866
ID ADS86866 standard; protein; 1755 AA.

XX AC ADS86866;

XX DT 16-DEC-2004 (first entry)

XX DE Human E3alpha ubiquitin ligase, hE3alphaII protein sequence.

XX KW Human; E3alpha ubiquitin ligase; huE3alphaII;

XX KW ubiquitin-proteasome pathway; chromosome 15 q; rapid muscle wasting;

XX KW fasting; metabolic acidosis; muscle degeneration; kidney failure;

XX KW renal cachexia; uremia; diabetes mellitus; sepsis; AIDS wasting syndrome;

XX KW cancer cachexia; negative nitrogen balance; burn; Cushing's syndrome;

XX KW inflammatory cachexia; hyperthyroidism; denervation atrophy;

XX KW protein/tissue wasting; energy-protein malnutrition; muscle atrophy;

XX KW gene therapy; enzyme.

XX OS Homo sapiens.

XX PN US2004185037-A1.

XX PD 23-SEP-2004.

XX PP 15-JAN-2004; 2004US-00758672.

XX PR 08-MAR-2000; 2000US-0187911P.

XX PR 28-NOV-2000; 2000US-00724126.

XX PA (HANH/) HAN H.

XX PA (KWAK/) KWAK K.

XX PI Han H, Kwak K;

XX WPI: 2004-707854/69.

XX N-PSDB; ADS86865.

XX PT Novel isolated human E3alpha ubiquitin ligase nucleic acid molecule

XX PT useful for treating and/or preventing renal cachexia or inflammatory

XX PT cachexia.

XX PS Claim 13; SEQ ID NO 4; 115pp; English.

XX CC The present invention relates to new orthologues of human E3alpha ubiquitin ligase, huE3alphaI and huE3alphaII. Most intracellular proteins are degraded through the ubiquitin-proteasome pathway. Proteins are marked for proteasomal degradation by conjugation of ubiquitin to the protein. Conjugation of the ubiquitin molecule involves the activation by E1 enzyme, subsequent transfer to E2 enzyme, which serves as a carrier, and then interacts with a specific E3 ligase family member. E3 ligase binds to proteins targeted for degradation and catalyzes the transfer of ubiquitin from the E2 carrier enzyme to the target protein. E3 ligase determines the specificity of the system. The E3alpha family is the main family of intracellular ligases and is involved in the N-end rule pathway of protein degradation. E3alpha enzyme binds directly to the primary destabilizing N-terminal amino acid and catalyzes ubiquitin conjugation thereby targeting the protein for degradation. The human E3alpha gene is located on chromosome 15 q. Increased proteolysis through the ubiquitin-proteasome pathway has been determined to be a major cause of rapid muscle wasting including, fasting, metabolic acidosis, muscle degeneration, kidney failure, renal cachexia, uremia, diabetes mellitus, sepsis, AIDS wasting syndrome, cancer cachexia, negative nitrogen balance, burns, Cushing's syndrome, inflammatory cachexia, hyperthyroidism, denervation atrophy, protein/tissue wasting, and energy-protein malnutrition. E3alpha plays a role in the overall increase in ubiquitination that is associated with and may mediate muscle atrophy in catabolic and other disease states. Treatment may be administered by gene therapy, cell therapy and antisense therapy methods. The present sequence is human E3alphaII protein sequence.

XX SQ Sequence 1755 AA;

XX PF 28-NOV-2000; 2000US-00724126.
 XX PR 08-MAR-2000; 2000US-0187911P.
 XX XX (AMGE-) AMGEN INC.
 XX PA Han H, Kwak K;
 XX PI WPI; 2004-236723/22.
 XX DR N-PSDB; ADJ95445.
 XX XX New nucleic acid molecule, useful for preparing a composition for
 XX PT diagnosing, treating or preventing diseases associated with human
 XX PT E3approximately polypeptide, e.g., muscle atrophy.
 XX XX Example 2; SEQ ID NO 6; 104pp; English.
 XX CC The invention relates to a new isolated nucleic acid molecule appearing
 XX CC as ADJ95441 (or its complement) encoding a ubiquitin ligase E3alpha 1
 XX CC protein appearing as ADJ95442. Also included are a vector comprising the
 XX CC nucleic acid, a host cell comprising the vector, a process of producing a
 XX CC E3alpha 1 ubiquitin ligase polypeptide, a composition comprising the
 XX CC nucleic acid molecule, a reagent comprising a detectably labelled
 XX CC nucleotide, and a method for determining the presence of a human E3alpha
 XX CC 1 ubiquitin ligase nucleic acid in a biological sample. The nucleic acid
 XX CC molecule is useful for preparing a composition for diagnosing, treating
 XX CC or preventing diseases associated with human E3alpha 1 polypeptide, e.g.
 XX CC muscle atrophy, cachexia, catabolic disorders, cancer cachexia, renal
 XX CC cachexia, inflammatory cachexia, muscle wasting disorders associated with
 XX CC metabolic acidosis, uremia, burns, hyperthyroidism, Cushing's syndrome,
 XX CC fasting, denervation atrophy, diabetes mellitus, sepsis and AIDS wasting
 XX CC syndrome. The present sequence represents mouse E3alpha 1.
 XX SQ Sequence 1755 AA;

Query Match 46.4%; Score 4278; DB 8; Length 1755;
 Best Local Similarity 46.5%; Pred. No. 0;
 Matches 835; Conservative 341; Mismatches 534; Indels 84; Gaps 28;

QY 1 MADE---EAGGTER--METSAELPQPPQRLASWQDQVPTAFLLHLAQLVPELYFAEM 55
 DB 1 MASEMEPFVQADRLSLECSAE-----EAGRWLQATDLNREVYQHLACHVFKYICRGP 54
 QY 56 DPDLKQESVQMSIFPLEWYLFGBDDPDICLEKLKHSQ-AFQCLGRVFKSGETTYSCRD 114
 DB 55 NPPFPQEDTLAHLILGPNWYICAEADPALGPPKLEQANKPSHLCLGRVFKSGEPTYSRD 114
 QY 115 CALDPTCVLCMDQFQDSVHNHRYKMTSTGGGFCDCGTEAWKTPPCVNHEPGRAGTI 174
 DB 115 CAVDPTCVLCMECLFGLSIHRDHYRMTTSGGGGFCDCGTEAWKTEGFCYQKHLSSRYV 174
 QY 175 -KENSRCPLNEEVIQARKIFPSVIKVVYEMTWESEKELPPELQIREKNERYCVLND 233
 DB 175 EEDPLVLHUSEVDIARTNYFALMFRYAVDIUTWEKESLPDELEVAESDITYCMFLND 234
 QY 234 EHSYDHVITYSLQALDCELABAQLHTTADKEGRRAVAGAYAAQEAKEBIDKHSENV 293
 DB 235 EVHTYEQVITYLQKAVNCYQKEAIGPATTVDRGRPRVYGFQYCDQAKTIVENTSRQ 294
 QY 294 SQHPLHVEVLHSEIMAHQKFAIRLGSWMNKIMSYSDSPFIQFCQACLRPEPSENPLIS 353
 DB 295 TK-PLKVQVHMSVAHQNFGLKALSWLGSYSDGLRILLCVQLQSGPDGENSESLVD 353
 QY 354 RLMLWDKLYKCAKTLHELIPSSPFMEYKGLPAMEPVKYKQLOKEYISDDHDSIS 413
 DB 354 RLMLNDSKLWKGARSYYHQLFSSSLMDLKYKLFALRFAKNYRQLQRFMEDDHERAVS 413
 QY 414 ITALSQMTPTVTLARHLEEQNVISVITETLELPEYLDNRNKNFQGS---QDKLG 470
 DB 414 VTALSQVQPTATLAEMLLTENMLTVIIKAFMDHL-KERDAQGRFOFERYTALQAFKR 472
 QY 471 RYVAVICDLKYLISKPTITWTERLRMQFLGFRSFLKILTCMQGMEIRRQVGHIEVDP 530

DB 473 RVQSLILDLKYVLISKRPTWSDLEKQFLQGGDALELLKCKMQGMDPITRQVQGHIEMP 532
 QY 531 DWEAAIAIQMLQNLILFQEWACACDELLLVAYKECHKAVMRCSTSFSSSKTVQVS-C 589
 DB 533 EWEAAFTLQMKLTHVISWQDWCALDEKVLIERAYKKCLAVLTOCHGFTDGRQPTLSIC 592
 QY 590 GHSLETKSYRVEDLVSILPLSRILAGLHVLSRLGAVSRHLHFVSPDFQVEVLVVEYP 649
 DB 593 GHSVETIRYCVSQEKSIVHLPISTRLAGLHVLSLSKSEVAYKPELLPLSELSPMLIEHP 652
 QY 650 LKCLVLVAQVAVEMRRNGLSLISOVFFYODVKCREMYDKDIIMLQICASIMDNKPTLL 709
 DB 653 LKCLVLCQVHAGMWRNGFSLVNGIYYTHNVKCRREMFDKDIIMLQTVGSMMDNHFLL 712
 QY 710 LVLYRYELAEAFN-----KTISTK--DQDLIKQYNTLIEMLQVLIYIVGERYVPGVNV 762
 DB 713 IMLSRFELVQLPSTPDYGRPSSEVTHKDVQNNTLIEMLYLIIMLVGERPNPGVGV 772
 QY 763 TKBEVTMBRIIHLICIEBPMPSIAIKNLPENNETGTLENINKVATFKKPGVSGHVYE 822
 DB 773 AATDEIKREITHQLSIKPMASSELVKSLPEDENKETGMEVIESVAHFKKPGLTGRGMYE 832
 QY 823 LKDESLKDPNMYFYHYSKTQHSKASHMOKRKRQENKDEALPPPPPPPCFAPSKVINLL 882
 DB 833 LKPECAKGFNLFFYHFSRABOSKABEQRKLRKREKDTALPPPPPPCLPFLASVLNL 892
 QY 883 NCDIMMYTILRTVFERAIDTDSNLTNEGMLQMAFHILALGLLEKQKQLOKAPSEEV-TFDF 941
 DB 893 QCDVMLYIMTILQWAVEHGSANSEMLORVHLHIGMALQEBKHLENAVEGHVQTFTE 952
 QY 942 YHKASRLGSSANN-----IQMLLEKLAGIPQLEQKQDMITWILQMDPTVKRLREKSLIVAT 998
 DB 953 TOKISKPGDAPHNSPSILAMLETQLQWAPSLAEAKDMIRWLLKWFNAIKKIRE--CSSSP 1010
 QY 999 TSGSESIKNDEITHDKEKAKRKAEEAARLHRQKIMQASALQKFIETHKLMYDNTSEM 1058
 DB 1011 VABEAGTMESSRDKDAERKRKKAIRLARREKIMQASEMRQHPIDENKELFOOTLEL 1070
 QY 1059 PGKEDSIMEESTPAVSYSRIALGPKRGPSVTEKEVLTCLCQEQEQVKIENNAVLISA 1118
 DB 1071 DTSASATL--DSSPPVSDAALTALGPAQTQVPEPQFVTCILCQEQEVTVGSRAVLAA 1128
 QY 1119 CVQKSTALTQHRKPIELSGEALDPLFMDPLAYGTYTSGCGHVMHVCWKYFRAVQUS 1178
 DB 1129 FYQSTVLSKORTKI-ADPEKYDPLFMHPDLSCGTHTSGCGHVMHACHQRYFDSVQAK 1187
 QY 1179 SQO-----RIHVDLPDESSEVLCPLCKSLCNTVPIIPLQPKINSEADALAQLLTIA 1233
 DB 1188 EQRROORLRLHTS-YDVENGEPCLPCBCLSNVTIPLL-LPPRSILSRILN-FSDOPDLA 1244
 QY 1234 RMIQTVLARISGNYI---RHAKGEN-----PIPIPFNQMGDSTLEFHSILSF 1278
 DB 1245 QWTRAVTQIKVQVQMLRKHNAADTSSESTEAMNIIPIPEGFRP-----DFY----- 1292
 QY 1279 GVSESIKYSNISKEMVILPATIYRIGLKVPPDERDPVPMLTWTSTCAFTIOAENLGD 1338
 DB 1293 ---PRNPYSDSIKEMLTFTGTAAYKVLKGVHNEGDPVPVILCWGTCAVTIOIERILD 1349
 QY 1339 EGKPLFGALQNRQHNGKALMQPAVAQRITCPQVLIQKHLVLLSVLNPISSEDTPLL 1398
 DB 1350 BEKPYFGPLPCRLDDCLSLRTRFAAAHWTVALLPVVQGHFCKLPSLPSDSYEDLPCLL 1409
 QY 1399 STDPLHVLVGAVLAPPSLYWDDPVDLQPSVSSSVSNHLYLPHLITMAHMLQILL---TVD 1455
 DB 1410 DIDMFHLVGLVLAFLPALQCDQ-----FSGSLATGDLHLFHLVTHAHLVQILLTSCTEE 1464
 QY 1456 TGLPLAQVQDESEEAHSASSPFAEISQYTSIGDCIDIPGWLWVSLKNGKITPYLCAALF 1515
 DB 1465 NGMD--QENPTGEELAILSLHKLHQYTSALKEAPSGHWRVSRVAAIMPFLKCSALF 1522
 QY 1516 FHYLLGVTPPELHTNSAGEYSALCSYLSLPTLPLFLFOEYDWTDTVRPLQWCAADPALL 1575

1523 PHYLVGPPAPDQLQ-SGTSHEPHLCNYLSLPTNLHLHQENSIMNSLIESWCQNSVK 1581
1576 NCLKOKNTVVRPRKRNLSIELPDDYSCLLNOASHFRCPRSADDERKHPVLCLFCALIC 1635
1582 RYLNGERGAIYPRGANKLIDLPEYSSLIQASNFSCPKSGDKSRAPTLCLVCGSLC 1641
1636 SONICCOETVNGEYGACIFHALHCGAGVIFPKIRECRVVLVEKGKARGCAYAPVLDREY 1695
1642 SUSYCCQAELEBEDVGACTAHTYSCSGNGIFLRVRCQVLFLAGTKGCFSPYLDY 1701
1696 GETDPLKRGNPLHLRSRYRKHLVWQOHCIEBIARSQETNQLFGFNWOLL 1749
1702 GETDQLRRGNPLHLCOERFRKIQLWQOHSITEIGHAQEANOQLVGIDWQHL 1755

RESULT 14
ADS86868
ID ADS86868 standard; protein; 1755 AA.
XX AC ADS86868;
XX DT 16-DEC-2004 (first entry)
XX DE Murine E3alpha ubiquitin ligase, muE3alphaII protein sequence.
XX KW Mouse; E3alpha ubiquitin ligase; muE3alphaII;
KW ubiquitin-proteasome pathway; rapid muscle wasting; fasting;
KW metabolic acidosis; muscle degeneration; kidney failure; renal cachexia;
KW uremia; diabetes mellitus; sepsis; AIDS wasting syndrome;
KW cancer cachexia; negative nitrogen balance; burn; Cushing's syndrome;
KW inflammatory cachexia; hyperthyroidism; denervation atrophy;
KW protein/tissue wasting; energy-protein malnutrition; muscle atrophy;
KW Gene therapy; enzyme.
XX OS Mus musculus.
XX PN US2004185037-A1.
XX DT 23-SEP-2004.
XX 15-JAN-2004; 2004US-00758672.
XX 08-MAR-2000; 2000US-0187911P.
XX 28-NOV-2000; 2000US-00724126.
XX (HANH/) HAN H.
XX PA (KWAK/) KWAK K.
XX PI Han H, Kwak K;
XX WPI; 2004-707854/69.
XX N-PSDB; ADS86867.
XX Novel isolated human E3alpha ubiquitin ligase nucleic acid molecule
XX useful for treating and/or preventing renal cachexia or inflammatory
XX cachexia.
XX Disclosure; SEQ ID NO 6; 115pp; English.
XX The present invention relates to new orthologs of human E3alpha
XX ubiquitin ligase, huE3alphaI and huE3alphaII. Most intracellular proteins
XX are degraded through the ubiquitin-proteasome pathway. Proteins are
XX marked for proteasomal degradation by conjugation of ubiquitin to the
XX protein. Conjugation of the ubiquitin molecule involves the activation by
XX E1 enzyme, subsequent transfer to E2 enzyme, which serves as a carrier,
XX and then interacts with a specific E3 ligase family member. E3 ligase
XX binds to proteins targeted for degradation and catalyses the transfer of
XX ubiquitin from the E2 carrier enzyme to the target protein. E3 ligase
XX determines the specificity of the system. The E3alpha family is the main
XX family of intracellular ligases and is involved in the N-end rule pathway
XX of protein degradation. E3alpha enzyme binds directly to the primary
XX destabilising N-terminal amino acid and catalyses ubiquitin conjugation
XX thereby targeting the protein for degradation. The human E3alpha gene is

CC located on chromosome 15 q. Increased proteolysis through the ubiquitin-
CC proteasome pathway has been determined to be a major cause of rapid
CC muscle wasting including, fasting, metabolic acidosis, muscle
CC degeneration, kidney failure, renal cachexia, uremia, diabetes mellitus,
CC sepsis, AIDS wasting syndrome, cancer cachexia, negative nitrogen
CC balance, burns, Cushing's syndrome, inflammatory cachexia,
CC hyperthyroidism, denervation atrophy, protein/tissue wasting, and energy-
CC protein malnutrition. E3alpha plays a role in the overall increase in
CC ubiquitination that is associated with and may mediate muscle atrophy in
CC catabolic and other disease states. Treatment may be administered by gene
CC therapy, cell therapy and antisense therapy methods. The present sequence
CC is the mouse E3alphaII protein sequence.
XX Sequence 1755 AA;
SQ

Query Match 46.4%; Score 4278; DB 8; Length 1755;
Best Local Similarity 46.5%; Pred. No. 0;
Matches 835; Conservative 341; Mismatches 534; Indels 84; Gaps 28;
QY 1 MADB---RAGGTER--MEISAEPLQTPORLASWMDQVDFYTAFLHLAQLVPEIYPAEM 55
DB 1 MASEPEPEVQAIKSLLECSAE-----EIAGRWLQATDLNREYVQHLAHCVPKLYCRGP 54
QY 56 DPDLKQESVQMSIFTPLEWYLFGEEDPDICLEKLKHSG-AFQLCGRVFKSGETTYSCRD 114
DB 55 NPPQKEDTLAQHILLGPMWYICAEDEPALGFPKLEQANKPSHLGCRVFKVGEPTYSCRD 114
QY 115 CAIDPTCVLCHMDQFQDSVHKHRRVKNMTSTGGGCDGCDTEAWTKGFCVNHPEGRAGTI 174
DB 115 CAVDPCTCVLMECFGLSITHRDHYRMTTSGGGGFCDCGDTAEWKEGPYCQKHLSSSSV 174
QY 175 -KENSRCPLNEVILVOARKIPPSVIVVEMTIVWEEKLELPELQIRKSNRYCYVLNPD 233
DB 175 EEDPLVHLSDEVIAARTNIFAIMFRYAVDILTWKESLEPDLVAKSSTIYICMLNPD 234
QY 234 EHSYDHYVYSLQALDCELAQAHLTTAIDKEGRVAVKAGAYACACBAKEDIKSHSENV 293
DB 235 EVHTYEVYITLQXAVNCTQKEAIGFATTVDRDGRRPVRYGDFQYCDQAKTVIRNTSRQ 294
QY 294 SOHPLHVEVLHSEIMAHOKPALRLGSMWNKMSYSSDFRQIFCQACLBEEDSDSENPLIS 353
DB 295 TK-PLKVQVMHSSVAAHQNFGLKALSGVIGYSDGLRLILCQVGLQEGDPGENSSLVD 353
QY 354 RLMLWDAKLYKGARKILHELIFSSFFMEMEKCLPAMEFVKYKQLOKEYISDDHDSIS 413
DB 354 RLMLNDSKLVGARSVHQLFMSSLLMDLTKYKLPALFPAKNYRQLQDFWEDDHERAVS 413
QY 414 ITALSVMQFTVPTLARHLIEBQNVISITETLLEVLPEYLDNRNKNFQGYV---ODKLG 470
DB 414 VTALSVMQFTVPTLARMLLTENLMTVIIKAPMDHL-KHRDAQGRFPERYTALQAKFR 472
QY 471 RYAVICDLKYLISKPTIWTTERLMQFLGFRSPFLKILTCMOGHEETRRQVGOHIEVDP 530
DB 473 RVQSLILDKVLISKPTSEWDELKQFLQGFDAFLLELLKCMQGMDFITRQVGOHIEBP 532
QY 531 DWBAIAIOMOLKNILLMFQEWACDBELLVAVYKECHKAVMRCSTSISSKTVQVS-C 589
DB 533 EWEAFTLQMKLTHVISVMQWCAUDEKVLTEAYKKCLAVITQCHGGFTDGEQPTLSIC 592
QY 590 GHSLETKSYRVEDLSVSIHLPLSRTLGLHVRSLRGAVSRHLHPEVSPEDFQVEVLVEYP 649
DB 593 GHSVETIRYCVSQBKVSIIHLPSIRLLAGLHLLSKSEVAYKFPPELLPLSELSPPLIEHP 652
QY 650 LRCLVLVAQVAVMWRNGLSLISQVFFYQDVCKEEMVDYDIIMLOIGASIMDNKELL 709
DB 653 LRCLVLCQVHAGMWRNRGFSLVNQIYYHYHNKCRREMFDDIVMLQGVSNMDFHFLM 712
QY 710 LVLORYELAEAFN-----KTISTK--DODLTKOYNTLIEEMLOVLIYVGERVYVGVNV 762
DB 713 IMLSRFELYQFSTPDYKRRFSSEVTHKDVVQOONTLIEEMLYLIIMLVGERFNFVGQV 772
QY 763 TKBEYTMREIHLCLIEPMHPSATAKNIPENNNETGLNVINKVATFKCPGVSGHGYVE 822

Db 773 AATDEIKREIHHQSLKPMASSELVSKLPEDENKTCMESVIESVAHFKKPGLTGRGMYE 832
QY 823 LKDESUKDFMYHYTHYSKTHQSKAEHMQKRRKQKQENKDBALPPPPPPPCAFSPKVINLL 882
Db 833 LKPECAKEFNLYTHYSRAEQSKAEARQLKREKEDTALPPALPPFCPLPASLVNLL 892
QY 883 NCDIMVYLTLTVERALDTSNLTGMLQMAFHILALGLERKQOLQKAPESV-TEDP 941
Db 893 QCDVLMYIMTIIQWAVEHGHGSAWSESMLQRLVHLIGMALQOEKHLLENAVEGHVQTF 952
QY 942 YHKASRLGGSAMN---IQMLELKLKGIPOLEGOKDMITWLOMPDVTVKRLREKSCLI 998
Db 953 TQIKSPGDAPHNSPSILAMLETIQNAPSLEAHKOMIRLLKMFNAKKIRE--CSSSSP 1010
QY 999 TSSSESINKDEITHDKAKRKKKAEAAARLHRQKIMQMSALQKNFTIETHKLMYDNTSEM 1058
Db 1011 VAEABGTMEESSRDKDKAKRKKAEATARLRREKINAQSEMQRHFDENKELPQOTLEL 1070
QY 1059 PGKEDSTMEESTPAVSYSRIALGPKRGPVSVEKEVLTCLCOBEOEVKIENNAVL 1118
Db 1071 DTSASATL--DSBPVPSDAALTALGPAQTQVPPRPQVPTCILCOBEOEVTGSRAMV 1128
QY 1119 CVQKSTALTQHRGKPTIELSGEALDPLFMDPLAVGTGSCGHVMAVCHQKYPFAVQLS 1178
Db 1129 FVORSTVLSKORTKI-ADPEKYDPLFMHFDLSCGTHGSCGHVMAHCHQWRPDSVQAK 1187
QY 1179 SQO-----RIHVDLFDLESGEYLCPLCKSLCNTVPIPIPOKQKINSENADALAL 1233
Db 1188 EQRQORLRLHTS-YDVENGELFCLCECLSNVTIPLL-LPPRSILSRRLN-FSDQPDLA 1244
QY 1234 RWTQTVLARISGNYI---RHAKEN-----PIPIFNQMGDSTLEFHSILSF 1278
Db 1245 QMTRAVTQQIKVQMLRKKNAAATGSSSETEAMNIPIPEGRFP-----DFY---- 1292
QY 1279 GVSESSIKSINKEMVLPATTYRIGLKVPPDERPRVPLMTSTCAPTIAIENLGD 1338
Db 1293 ---PRPYSISIKEMLTFTGTAAKYGLKVHPNEGDRPVLCWCATYIQSRIELSD 1349
QY 1339 EGKPLFGALQNRHNGIKALMQFAVAQRTICPQVLQKHLVRLSVLPNIKESDTPCLL 1398
Db 1350 EEPVFGPLPCRLDDCLSLRTPAAAHWTVALPVVQGHFCKLPASIVPSDSVEDLPCIL 1409
QY 1399 SIDLFVVLGAVLAPSLYWDPDVLOPSVSSSYNNHLYLPHLITWAMLMQIILL----TVD 1455
Db 1410 DIDMFHLLVGLVLAFFALQOCD-----FSGSSLATGDLHIFHLVTHAHIVQIILLT 1464
QY 1456 TGLPLAQVQDSSEAHSSAFFAEISQYTSIGSGCDIPGVLVSLKNGITPYLRCAALF 1515
Db 1465 NGMD--QENPTGEEELAILSLHTLHQYTGSALKEAPSGHMLWRSVRAAIMPFLKCALF 1522
QY 1516 FHYLLGVTPPEELHTNSAEGEYALCSYLSLPTNLFLFOEYWDTVRPLLRWCADPALL 1575
Db 1523 FHYLVGVAPPDQLQV-SGTHPEHLNLYLSLPTNLHLFOENSIMNSLIESWCQNSEVK 1581
QY 1576 NCLKQKNTVVRPRKNSLIHELDDYSCLLNQASHFRCPASADDERKHVPLCLFCGAILC 1635
Db 1582 RYLNGERGATSYPRGANKLJLDPEYSSLIQNASFNCPKSGGDKRAPTLCLVCGSLLC 1641
QY 1636 SONICCOEIVNGEVGACIFALHCGAGVCIELKIRECRVVLVEGRGCAYPAPYLDEY 1695
Db 1642 SOSYCCOALEGSDVGACTAHTYSCSGGAGIFLRVRECQVFLAGTKGCFYSPYLLDDY 1701
QY 1696 GETDPGLKRNPNLHLRERVKRLHLVWQHCIIIEIARSQETNQMLPGFNWQLL 1749
Db 1702 GETDOGLRGNPNLHLCOERFKYQKLMQOHSITEEIGHAQAQTLVGIDWQHL 1755

RESULT 15

AAB93464

ID AAB93464 standard; protein; 811 AA.

XX

AC AAB93464;

XX

DT 26-JUN-2001 (first entry)

XX Human protein sequence SEQ ID NO:12732.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.

XX Homo sapiens.

XX EP1074617-A2.

XX 07-FEB-2001.

XX 28-JUL-2000; 2000EP-00116126.

XX 29-JUL-1999; 99JP-00248036.

XX 27-AUG-1999; 99JP-00300253.

XX 11-JAN-2000; 2000JP-00118776.

XX 02-MAY-2000; 2000JP-00183767.

XX 09-JUN-2000; 2000JP-00241899.

XX (HELI-) HELIX RES INST.

XX Ota T, Isogai T, Nishikawa T, Hayaashi K, Saito K, Yamamoto J;

XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.

XX Claim 8; SEQ ID NO 12732; 2537pp + Sequence Listing; English.

XX The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dr primer and an oligonucleotide complementary to the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention

XX Sequence 811 AA;

Query Match 45.6%; Score 4209; DB 4; Length 811;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 811; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 204 MTIWEKEKLPPELOIREKNERYCYLVFNDEHHSYDVHYSYVLSQALDCELAQAHLTTAI 263

Db 1 MTIWEKEKLPPELOIREKNERYCYLVFNDEHHSYDVHYSYVLSQALDCELAQAHLTTAI 60

QY 264 DKEGRAVKAAGAACQAEKEDIKSHSENVSHPLHVEVLHSEIMAHQKFAFLRGSMWNK 323

Db 61 DKEGRAVKAAGAACQAEKEDIKSHSENVSHPLHVEVLHSEIMAHQKFAFLRGSMWNK 120

QY 324 IMSYSSDFQIFQACLRPEPSPENCLISRLMLWDAKLYKGARKILHILFSPFMEME 303

```

Db 121 IMSYSDFRQIFQACLRBPDPSENCLISRLMLWDKLYKGARKIILHELIFSPFME 180
Qy 384 YKLFAMEFYKYKQLOKEYISDDHRSISITALSVMFTVPTLARHLIEQNVISVITE 443
Db 181 YKLFAMEFYKYKQLOKEYISDDHRSISITALSVMFTVPTLARHLIEQNVISVITE 240
Qy 444 TLEVLPEYLDNRNKNFQGYSDQKLRVYAVICDLKYILISKPTIWTIERLMOFLEGFR 503
Db 241 TLEVLPEYLDNRNKNFQGYSDQKLRVYAVICDLKYILISKPTIWTIERLMOFLEGFR 300
Qy 504 SFLKILTCMQMBEIRQVQCHIEVDPDWEAAIAIQMLXNIIILMFQEWACDELLVA 563
Db 301 SFLKILTCMQMBEIRQVQCHIEVDPDWEAAIAIQMLXNIIILMFQEWACDELLVA 360
Qy 564 YKECHKAVMRCSTSFISSTVTVQSCGHSLETYSYRVEDLVSIHPLSRTLAGLHVRLS 623
Db 361 YKECHKAVMRCSTSFISSTVTVQSCGHSLETYSYRVEDLVSIHPLSRTLAGLHVRLS 420
Qy 624 RLGAVERLHEFVSFEDFQVEVLVEYPLRCLVLVAQVVAEMWRRNGLSLSQVFFYQDVKC 683
Db 421 RLGAVERLHEFVSFEDFQVEVLVEYPLRCLVLVAQVVAEMWRRNGLSLSQVFFYQDVKC 480
Qy 684 REEMYDKDIIIMLQIGASLMDPNKFLLLVLQRYELAEAFNKTIISTKQODLIKQYNTLIEEM 743
Db 481 REEMYDKDIIIMLQIGASLMDPNKFLLLVLQRYELAEAFNKTIISTKQODLIKQYNTLIEEM 540
Qy 744 LQVLIYIVGERYVPGVGNVTKBEVTWREIILHLCIEBPMPSAIAKNLPENNETGLENV 803
Db 541 LQVLIYIVGERYVPGVGNVTKBEVTWREIILHLCIEBPMPSAIAKNLPENNETGLENV 600
Qy 804 INKVATFKPGVSGHGVYELKDSLDKDFNMVYHYSTQHSKAEHMOKKRRKOENKDEAL 863
Db 601 INKVATFKPGVSGHGVYELKDSLDKDFNMVYHYSTQHSKAEHMOKKRRKOENKDEAL 660
Qy 864 PPPPPPEFCFAPSKVINLLNCDIMMYILRTVFERAIDTDSNLWTEGMLQWAPHILALGLL 923
Db 661 PPPPPPEFCFAPSKVINLLNCDIMMYILRTVFERAIDTDSNLWTEGMLQWAPHILALGLL 720
Qy 924 EEKQOLOKAPEEVTFDFYHKASRLGSSAMNIOHLEKLGIPOLGQKDMITWILQMPD 983
Db 721 EEKQOLOKAPEEVTFDFYHKASRLGSSAMNIOHLEKLGIPOLGQKDMITWILQMPD 780
Qy 984 TVKRLREKSLIVATTSGSESINKDEITHDK 1014
Db 781 TVKRLREKSLIVATTSGSESINKDEITHDK 811

```

Search completed: July 9, 2005, 15:28:52
Job time : 141 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2005, 14:29:13 ; Search time 49 Seconds
(without alignments)
2664.515 Million cell updates/sec

Title: US-10-758-672A-2
Perfect score: 9224
Sequence: 1 MADEAGGTERMEISAEPLPQ.....BIARQETNQLFGFNWQLL 1749

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9224	100.0	1749	4	US-09-724-126A-2
2	9087	98.5	1734	4	US-09-724-126A-19
3	8617	93.4	1757	4	US-09-724-126A-15
4	4327	46.9	1755	4	US-09-724-126A-4
5	4278	46.4	1755	4	US-09-724-126A-6
6	456	4.9	223	4	US-09-270-767-46327
7	450.5	4.9	334	4	US-09-270-767-45019
8	317.5	3.4	817	4	US-09-248-796A-19402
9	274.5	3.0	631	4	US-09-248-796A-19403
10	197	2.1	280	4	US-09-270-767-32482
11	190.5	2.1	3259	4	US-09-949-016-6507
12	184	2.0	1979	4	US-09-949-016-6468
13	184	2.0	2047	4	US-09-949-016-7404
14	176	1.9	2662	4	US-09-595-684B-31
15	175.5	1.9	1427	4	US-09-538-092-1044
16	175.5	1.9	2663	4	US-09-538-092-1252
17	172.5	1.9	1780	4	US-09-949-016-6899
18	172.5	1.9	1786	4	US-09-949-016-7880
19	172	1.9	1055	4	US-09-949-016-9776
20	172	1.9	3878	4	US-09-514-253-11
21	171	1.9	994	4	US-09-949-016-6779
22	164	1.8	2954	4	US-09-150-867-1
23	159.5	1.7	2733	4	US-09-949-016-11433
24	157.5	1.7	2789	4	US-09-949-016-8208
25	156.5	1.7	2710	1	US-08-480-604A-6
26	156.5	1.7	2710	2	US-08-405-496A-6
27	156.5	1.7	2710	3	US-08-915-136-6

28	156.5	1.7	2710	3	US-08-957-310-6	Sequence 6, Appli
29	156.5	1.7	2710	4	US-10-011-366-6	Sequence 6, Appli
30	156.5	1.7	2710	4	US-09-084-517-6	Sequence 6, Appli
31	155.5	1.7	2704	4	US-09-538-092-1260	Sequence 1260, Ap
32	152.5	1.7	3433	4	US-09-538-092-1136	Sequence 1136, Ap
33	150.5	1.6	1388	3	US-09-723-262-2	Sequence 2, Appli
34	150.5	1.6	1388	3	US-09-723-262-2	Sequence 2, Appli
35	150.5	1.6	1388	3	US-09-723-262-2	Sequence 2, Appli
36	149	1.6	3433	4	US-09-091-501B-10	Sequence 10, Appl
37	148	1.6	2125	4	US-09-919-172-29	Sequence 29, Appl
38	148	1.6	2871	4	US-09-538-092-936	Sequence 936, App
39	146.5	1.6	973	4	US-09-392-714-24	Sequence 24, Appl
40	146	1.6	958	4	US-09-171-991-6	Sequence 6, Appli
41	145.5	1.6	1401	4	US-09-750-590A-2	Sequence 2, Appli
42	144.5	1.6	1530	4	US-09-949-016-6668	Sequence 6668, Ap
43	142.5	1.5	1362	4	US-09-949-016-7033	Sequence 7033, App
44	142.5	1.5	1530	4	US-09-976-594-736	Sequence 736, App
45	142.5	1.5	3210	4	US-09-538-092-1154	Sequence 1154, Ap

ALIGNMENTS

RESULT 1
US-09-724-126A-2
; Sequence 2, Application US/09724126A
; Patent No. 6706505
; GENERAL INFORMATION:
; APPLICANT: Han, Hui-Quan
; APPLICANT: Kwak, Keith
; TITLE OF INVENTION: Human E3 Alpha Ubiquitin Ligase Family
; FILE REFERENCE: 01017735966A
; CURRENT APPLICATION NUMBER: US/09/724,126A
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 60/187,211
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 2
; LENGTH: 1749
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-126A-2

Query Match	100.0%	Score 9224;	DB 4;	Length 1749;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1749;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MADBEAGGTERMEISAEPLPQTPQRLASWWDQVDTAFLHHLAQLVPEIYPAEMDPDL 60		
Db	1	MADBEAGGTERMEISAEPLPQTPQRLASWWDQVDTAFLHHLAQLVPEIYPAEMDPDL 60		
QY	61	KOESVQMSIETPLEWYLFSGDDPDICLEKXHGAFQICGRVFKSGETTYSCRDCAIDPT 120		
Db	61	KOESVQMSIETPLEWYLFSGDDPDICLEKXHGAFQICGRVFKSGETTYSCRDCAIDPT 120		
QY	121	CVLCMDCFQDSVHKHRYKMTSTGGFCDCGDTAWKTGFCVNHHPGRAGTIKENSRC 180		
Db	121	CVLCMDCFQDSVHKHRYKMTSTGGFCDCGDTAWKTGFCVNHHPGRAGTIKENSRC 180		
QY	181	FLNEBIVQARKIPPSVIKYVEMTIMBEKELPELOIRKERNERYCVLFNDEHHSYDH 240		
Db	181	FLNEBIVQARKIPPSVIKYVEMTIMBEKELPELOIRKERNERYCVLFNDEHHSYDH 240		
QY	241	VIYSLQALDCELAQAQLHTTAIDKEGRRAYKAGYACQAKEDIKSHSENVSHPLHV 300		
Db	241	VIYSLQALDCELAQAQLHTTAIDKEGRRAYKAGYACQAKEDIKSHSENVSHPLHV 300		
QY	301	EVLHSEIMAHQKFAIRLGSWNKMISYSSDFQICQACLRPEEPPSENPCILSRMLWDA 360		
Db	301	EVLHSEIMAHQKFAIRLGSWNKMISYSSDFQICQACLRPEEPPSENPCILSRMLWDA 360		
QY	361	KLYGARKILHELIFSSFFMEMEYKCLFAMBFVYKQLOKEYISDDHRSISITALSVDQ 420		

Db 361 KUYGARKILHBLIIFSSPFMEYKCLFAMEFVKYKQLQKEYISDDHRSISITALSVQ 420
QY 421 MFTVPTLARHLIEEQNVISVITETLLEVLPEYLDNRNKNFPGYSQDKLGRVYAVICDLK 480
Db 421 MFTVPTLARHLIEEQNVISVITETLLEVLPEYLDNRNKNFPGYSQDKLGRVYAVICDLK 480
QY 481 YLISKPTIWTBRLMQFLGFRSFLKILTCHQGHEEIRROVGOHIEVDPDWEAAIAIOM 540
Db 481 YLISKPTIWTBRLMQFLGFRSFLKILTCHQGHEEIRROVGOHIEVDPDWEAAIAIOM 540
QY 541 OLKNILLMFOEWACDEBELLVAYKECHKAVNRCSTSPISSSKTVVQSCGHSLETKSVRV 600
Db 541 OLKNILLMFOEWACDEBELLVAYKECHKAVNRCSTSPISSSKTVVQSCGHSLETKSVRV 600
QY 601 SEDLVSIHPLSRTLAGLHVRLSRGAVSRLEHFEVSPEDQVEVLVEYPLRCLVLVAQVV 660
Db 601 SEDLVSIHPLSRTLAGLHVRLSRGAVSRLEHFEVSPEDQVEVLVEYPLRCLVLVAQVV 660
QY 661 AEMWRNGLSLISQVFFYQDVKREMYDKOIMLQIGASLMDPNKFLLLVLQRYELABA 720
Db 661 AEMWRNGLSLISQVFFYQDVKREMYDKOIMLQIGASLMDPNKFLLLVLQRYELABA 720
QY 721 FNKTISTDODLIIKOYNTLIEBMLQVLIYVGERVPGVGNVTKBEVTRREIHHLLCIEP 780
Db 721 FNKTISTDODLIIKOYNTLIEBMLQVLIYVGERVPGVGNVTKBEVTRREIHHLLCIEP 780
QY 781 MPHSAIAKLPENNETGLENVINKVATFKKPGVSGHGVYELKDESLODFNMFYHYSK 840
Db 781 MPHSAIAKLPENNETGLENVINKVATFKKPGVSGHGVYELKDESLODFNMFYHYSK 840
QY 841 TQHSKAEHMQKRRKOENKDEALPPPPPPPCPAKSKVINLNCIDIMMYILRTVPERAID 900
Db 841 TQHSKAEHMQKRRKOENKDEALPPPPPPPCPAKSKVINLNCIDIMMYILRTVPERAID 900
QY 901 TDSNLTWTEGMLQMAFHILALGLEEKQLOKAPPEEVTDFVHKASRLGSSAMNLOMLE 960
Db 901 TDSNLTWTEGMLQMAFHILALGLEEKQLOKAPPEEVTDFVHKASRLGSSAMNLOMLE 960
QY 961 KLKGPQLSGQDMITWILQMFDTVKRLREKSLIVATTSGSEIKNDBITHDKEKAERK 1020
Db 961 KLKGPQLSGQDMITWILQMFDTVKRLREKSLIVATTSGSEIKNDBITHDKEKAERK 1020
QY 1021 RKAEBARLHROKIMQMSALQKNFIETHKLMDNTSEMPCKEDSIMEESTPAVSDYSRI 1080
Db 1021 RKAEBARLHROKIMQMSALQKNFIETHKLMDNTSEMPCKEDSIMEESTPAVSDYSRI 1080
QY 1081 ALGPKRGPSVTEKEVLTCLCOBEOBKLENNAMVLSACVQKSTALTQHRGKPIELSGEA 1140
Db 1081 ALGPKRGPSVTEKEVLTCLCOBEOBKLENNAMVLSACVQKSTALTQHRGKPIELSGEA 1140
QY 1141 LDPLFMDPLAYGTYTSGCGHVMHAYCWQKYFEAVQLSSQQRHIVDLPDLESGEYLCPIC 1200
Db 1141 LDPLFMDPLAYGTYTSGCGHVMHAYCWQKYFEAVQLSSQQRHIVDLPDLESGEYLCPIC 1200
QY 1201 KSLCNTVPIIPILOPKIMSENADALQLLTARWITQTVLARIISGNIRHAKGENPIPIF 1260
Db 1201 KSLCNTVPIIPILOPKIMSENADALQLLTARWITQTVLARIISGNIRHAKGENPIPIF 1260
QY 1261 FNQGMGDSLTLEPHSILSFGVSESIKYSNSIKEMVILFATTIYRIGLKVPPDERDPRVPM 1320
Db 1261 FNQGMGDSLTLEPHSILSFGVSESIKYSNSIKEMVILFATTIYRIGLKVPPDERDPRVPM 1320
QY 1321 TWSTCAFTTQAIENLLGDEGKPLFGALQNRQHNGLKALMQFAVAQRITCPQVLIQKHLVR 1380
Db 1321 TWSTCAFTTQAIENLLGDEGKPLFGALQNRQHNGLKALMQFAVAQRITCPQVLIQKHLVR 1380
QY 1381 LLSVLPNLIKSEDTPCLLSIDLPHVLVGAVLAPPPLYWDDPVDLQPSVSSSSNHLHYLFH 1440
Db 1381 LLSVLPNLIKSEDTPCLLSIDLPHVLVGAVLAPPPLYWDDPVDLQPSVSSSSNHLHYLFH 1440
QY 1441 LITMAHMLQILLTVDVTGLPLAQVQEDSEAHSSAFPAEISQYTSIGSGCDIPGWYLWVS 1500

Db 1441 LITMAHMLQILLTVDVTGLPLAQVQEDSEAHSSAFPAEISQYTSIGSGCDIPGWYLWVS 1500
QY 1501 LKNGITPYLRCAALFFHYLLGVTPPEELHTNSAGEYSALCSYLSLPTNLFLFQEYWD 1560
Db 1501 LKNGITPYLRCAALFFHYLLGVTPPEELHTNSAGEYSALCSYLSLPTNLFLFQEYWD 1560
QY 1561 VRPLLQWCAADPALINCLKOKNTVVRPRKNSLIELPDDYSCLINQASHFRCPRSADDE 1620
Db 1561 VRPLLQWCAADPALINCLKOKNTVVRPRKNSLIELPDDYSCLINQASHFRCPRSADDE 1620
QY 1621 RKHPVLCFLCGAIIICSONICCOEIVNGEVEGACIFPHALHCGAGVCIFLIKRECRVVLVEG 1680
Db 1621 RKHPVLCFLCGAIIICSONICCOEIVNGEVEGACIFPHALHCGAGVCIFLIKRECRVVLVEG 1680
QY 1681 KARGCAYPAPYLDYGETDPCGLKEGNPLHLSREYRKLHLVMOQHCIIIEETARSOETNOM 1740
Db 1681 KARGCAYPAPYLDYGETDPCGLKEGNPLHLSREYRKLHLVMOQHCIIIEETARSOETNOM 1740
QY 1741 LFGFNWOLL 1749
Db 1741 LFGFNWOLL 1749

RESULT 2
US-09-724-126A-19
; Sequence 19, Application US/09724126A
; Patent No. 6706505
; GENERAL INFORMATION:
; APPLICANT: Han, Hui-Quan
; APPLICANT: Kwak, Keith
; TITLE OF INVENTION: Human E3 Alpha Ubiquitin Ligase Family
; FILE REFERENCE: 01017/35966A
; CURRENT APPLICATION NUMBER: US/09/724,126A
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 60/187,211
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 19
; LENGTH: 1734
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-126A-19

Query Match 98.5%; Score 9087; DB 4; Length 1734;
Best Local Similarity 98.8%; Fred. No. 0;
Matches 1728; Conservative 0; Mismatches 1; Indels 20; Gaps 1;

QY 1 MADEAGGTERMEISAEPLPOTPQRLASWWDQVDPYTAFLHHLAQLVPEIYPAEMDPDLE 60
Db 6 MADEAGGTERMEISAEPLPOTPQRLASWWDQVDPYTAFLHHLAQLVPEIYPAEMDPDLE 65
QY 61 KOEESVQMSIPTPLEWYLFGEEDPDICLEKLKHSGAFQICGRVFKSGETTYSRCDCAIDPT 120
Db 66 KOEESVQMSIPTPLEWYLFGEEDPDICLEKLKHSGAFQICGRVFKSGETTYSRCDCAIDPT 125
QY 121 CVLCHMDCFQDSVHKHRYKMTSTGGFCDCGDTBAWTKGPCVNHPEGRAGTIKENSR 180
Db 126 CVLCHMDCFQDSVHKHRYKMTSTGGFCDCGDTBAWTKGPCVNHPEGRAGTIKENSR 185
QY 181 PLNEBVIIVQARKIPPSVIKYVEMVTIWEKEKELPPELOIREKNERYCYVLFNDEHSHVDH 240
Db 186 PLNEBVIIVQARKIPPSVIKYVEMVTIWEKEKELPPELOIREKNERYCYVLFNDEHSHVDH 245
QY 241 VIYSIQRALDCELAQAQLHTTAIDKEGRRAVKAGAYAAQCEAKEDIKSHSENVSHPLHV 300
Db 246 VIYSIQRALDCELAQAQLHTTAIDKEGRRAVKAGAYAAQCEAKEDIKSHSENVSHPLHV 305
QY 301 EYVLSHEIMAHQKFAIRLCSWMNKINSYSDRQIFQACLRPEEPSENPCILSRMLWDA 360
Db 306 EYVLSHEIMAHQKFAIRLCSWMNKINSYSDRQIFQACLRPEEPSENPCILSRMLWDA 365
QY 361 KLYKGARKILHBLIFSSPFMEYKCLFAMEFVKYKQLQKEYISDDHRSISITALSVQ 420

Db 366 KLYGARKILHELIFSSFFMEYKCLFAMFVKYKQKQKEYISDDHRSISITALSQ 425
Qy 421 MFTVPTLARHLIEQNVSIVITETLLEVLPEYLDNRNKNFQYSDQKLGRIYAVICDLK 480
Db 426 MFTVPTLARHLIEQNVSIVITETLLEVLPEYLDNRNKNFQYSDQKLGRIYAVICDLK 485
Qy 481 YLISKPTIWTBRLMOPLEGPSFLKILTCQGHIEIRROVQOHLIEVDPDWEAAIAIQM 540
Db 486 YLISKPTIWTBRLMOPLEGPSFLKILTCQGHIEIRROVQOHLIEVDPDWEAAIAIQM 545
Qy 541 QLNKILLMFOEWACDEEILLVAYKECHKAVNRCTSFISSTKVVSQSGHSLKTSYRV 600
Db 546 QLNKILLMFOEWACDEEILLVAYKECHKAVNRCTSFISSTKVVSQSGHSLKTSYRV 605
Qy 601 SEDLSIHLPLSRTLAGLHVRLSRGAVSRLEHFPVSFDFQVEVLVEYPLRCLVLVAQV 660
Db 606 SEDLSIHLPLSRTLAGLHVRLSRGAVSRLEHFPVSFDFQVEVLVEYPLRCLVLVAQV 665
Qy 661 AEWWRNGLSLSQVFFYODVKCRBMYDKDIIIMQIGASLMDPNKFLLLVLRVELAE 720
Db 666 AEWWRNGLSLSQVFFYODVKCRBMYDKDIIIMQIGASLMDPNKFLLLVLRVELAE 725
Qy 721 FNKTISTKQODLIIKOYNTLIEBMLQVLIYIGERVVPGVGNVTKBEVTWREIHLICIEP 780
Db 726 FNKTISTKQODLIIKOYNTLIEBMLQVLIYIGERVVPGVGNVTKBEVTWREIHLICIEP 785
Qy 781 MPHSAIAKNLPENNETGLENVINKVATFKPGVSGHGVYELKDBSLKDFNMYFYYSK 840
Db 786 MPHSAIAKNLPENNETGLENVINKVATFKPGVSGHGVYELKDBSLKDFNMYFYYSK 845
Qy 841 TQHSKAEHMKQKRGQENKDEALPPPPPEFCFAPSKVINLNCDIMMYILTVFERAID 900
Db 846 TQHSKAEHMKQKRGQENKDEALPPPPPEFCFAPSKVINLNCDIMMYILTVFERAID 905
Qy 901 TDSNLWTEGMLQMAHIALGLESQLOKQAPBEEVTFDFYHKASRLGSSAMNQLMLE 960
Db 906 TDSNLWTEGMLQMAHIALGLESQLOKQAPBEEVTFDFYHKASRLGSSAMNQLMLE 965
Qy 961 KLKGIPOLEGQKDMITWILQMFDTVKRLREKSLIVATTSGSEIKNDEITHKEKARK 1020
Db 966 KLKGIPOLEGQKDMITWILQMFDTVKRLREKSLIVATTSGSEIKNDEITHKEKARK 1025
Qy 1021 RBAEAAHLRQKIMQMSALQKQNFIEHKLMDYNTSEMPGKEDSIMEBESTPAVSDYSRI 1080
Db 1026 RBAEAAHLRQKIMQMSALQKQNFIEHKLMDYNTSEMPGKEDSIMEBESTPAVSDYSRI 1085
Qy 1081 ALGPKRGPSVTEKEVLTCILCOEEOVKIENNAMVLSACVOKSTALTQHRGKPIELSGEA 1140
Db 1086 ALGPKRGPSVTEKEVLTCILCOEEOVKIENNAMVLSACVOKSTALTQHRGKPIELSGEA 1145
Qy 1141 LDPLFMDPLAYGTYTSGCGHVMHACVQKYPEAVQLSSQQRHVDLPDLSGEYLCLPLC 1200
Db 1146 LDPLFMDPLAYGTYTSGCGHVMHACVQKYPEAVQLSSQQRHVDLPDLSGEYLCLPLC 1205
Qy 1201 KSLCNTVPIIPILOPKINSENADALAQLTLARWIQTVLARISGYNIHAKGENPIPIF 1260
Db 1206 KSLCNTVPIIPILOPKINSENADALAQLTLARWIQTVLARISGYNIHAKGENPIPIF 1265
Qy 1261 FNQMGDSTLEPHSILSRGVSSIKYSIKEMVILPATIIVIRIGLKYPPDPRVPML 1320
Db 1266 FNQMGDSTLEPHSILSRGVSSIKYSIKEMVILPATIIVIRIGLKYPPDPRVPML 1325
Qy 1321 TWSTCAFTIQTALENLIGDEGKLPFALQNRQNGKALKMQPVAQRITCPQVLIQKHLVR 1380
Db 1326 TWSTCAFTIQTALENLIGDEGKLPFALQNRQNGKALKMQPVAQRITCPQVLIQKHLVR 1385
Qy 1381 LLSVVLPNIKSBDTFCLLSIDLPHVLVGAVLAPPSLYWDDPVDLPQSSVSSSYNHLFLH 1440
Db 1386 LLSVVLPNIKSBDTFCLLSIDLPHVLVGAVLAPPSLYWDDPVDLPQSSVSSSYNHLFLH 1445
Qy 1441 LITMAHMLQILLTVDTGLPLAQVQDSEAHNSASSFFAISIQYTSSTGSCDIPGHWLWS 1500

Db 1446 LITMAHMLQILLTVDTGLPLAQVQDSEAHNSASSFFAISIQYTSSTGSCDIPGHWLWS 1505
Qy 1501 LKNGITPYLRCAALFFHYLLGVTTPPELHTNSABEYSALCSYLSLPTNLFLFLFOEYMDT 1560
Db 1506 LKNGITPYLRCAALFFHYLLGVTTPPELHTNSABEYSALCSYLSLPTNLFLFLFOEYMDT 1565
Qy 1561 VRPLLRCAADPALNCLKOKNTVVRYPKRNLSLIELPDDYSCLLNOASHFPCPSADDE 1620
Db 1566 VRPLLRCAADPALNCLKOKNTVVRYPKRNLSLIELPDDYSCLLNOASHFPCPSADDE 1625
Qy 1621 RHGPVLCIFCGAILCSQNICQOEIVNGBEVGACIFALHCCGAGVCIFLKIRECRVVLVEG 1680
Db 1626 RHGPVLCIFCGAILCSQNICQOEIVNGBEVGACIFALHCCGAGVCIFLKIRECRVVLVEG 1685
Qy 1681 KARGCAYPAPYLDEYGETDPGLKGNPLHLRSRVRKHLVWQQHCCIIEBIARSQETNOM 1740
Db 1686 KARGCAYPAPYLDEYGETDPGLKGNPLHLRSRVRKHLVWQQHCCIIEBIARSQETNOM 1745
Qy 1741 LFGFNWQLL 1749
Db 1726 LFGFNWQLL 1734

RESULT 3

US-09-724-126A-15

; Sequence 15, Application US/09724126A

; Patent No. 6706505

; GENERAL INFORMATION:

; APPLICANT: Han, Hui-Quan

; APPLICANT: Kwak, Keich

; TITLE OF INVENTION: Human E3 Alpha Ubiquitin Ligase Family

; FILE REFERENCE: 01077/35966A

; CURRENT APPLICATION NUMBER: US/09/724,126A

; CURRENT FILING DATE: 2000-11-28

; PRIOR APPLICATION NUMBER: US 60/187,211

; PRIOR FILING DATE: 1999-03-01

; NUMBER OF SEQ ID NOS: 29

; SOFTWARE: Patent in version 3.0

; SEQ ID NO 15

; LENGTH: 1757

; TYPE: PRT

; ORGANISM: Mouse

US-09-724-126A-15

Query Match 93.4%; Score 8617; DB 4; Length 1757;

Best Local Similarity 92.1%; Pred. No. 0;

Matches 1619; Conservative 64; Mismatches 66; Indels 8; Gaps 3;

Qy 1 MADEAGGTERMBISAEPLPQTPORLASWWDQVDFYTAFLHHLAQLVPEIYFAEMDPDLE 60
Db 1 MADEEMDGAERMDVSPPEPLAPQRPASWWDQVDFYTAFLHHLAQLVPEIYFAEMDPDLE 60
Qy 61 KOESVQMSIITPLEWYLFGEEDPDICLKLKXSGAFOLCGRVKSGETTYSCRDCAIDPT 120
Db 61 KOESVQMSIITPLEWYLFGEEDPDICLKLKXSGAFOLCGRVKSGETTYSCRDCAIDPT 120
Qy 121 CVLQMDCFQSDVHKHNRHYKMTSTGGGFCDCGDETEAWKTPFCVNHPEGRAGTICKENSR 180
Db 121 CVLQMDCFQSDVHKHNRHYKMTSTGGGFCDCGDETEAWKTPFCVNHPEGRAGTICKENSR 180
Qy 181 PLNBEVIVQARKIPSVIKYVEMTWIWEBEKELEPELOIRKERNERYCVLFNDEHSHSDH 240
Db 181 PLNBEVIVQARKIPSVIKYVEMTWIWEBEKELEPELOIRKERNERYCVLFNDEHSHSDH 240
Qy 241 VIYSQALDCELAEOHLHTTAIDKEGPRVAKAGYACQAKEDIKSHSENVSOHPLHV 300
Db 241 VIYSQALDCELAEOHLHTTAIDKEGPRVAKAGYACQAKEDIKSHSENVSOHPLHV 300
Qy 301 EVLHSEIMAHQKFAIRLGSMWNKIMSYSSDFQIFCOACLEEDPSENPCILSRMLMWA 360
Db 301 EVLHSEIMAHQKFAIRLGSMWNKIMSYSSDFQIFCOACLEEDPSENPCILSRMLMWA 360
Qy 361 KLYGARKILHELIFSSFFMEYKCLFAMFVKYKQKQKEYISDDHRSISITALSQ 420

Db 361 KLYKGARKLHLLHSSPFMEYKGLFAMEFVKYKQLOKEYISDDHERSISITALSQ 420
Qy 421 MFTVPTLARHLIEBQNVISVITETLLEVLPEYLDNNKFNFOGYSQDKLGRVYAVICDLK 480
Db 421 MLTVPPTLARHLIEBQNVISVITETLLEVLPEYLDNNKFNFOGYSQDKLGRVYAVICDLK 480
Qy 481 YLILSKPTTWTERLMOFLGFRSFLKILTCMOGHEETIRROGHLEVDPDMEAAIAIQM 540
Db 481 YLILSKPTTWTERLMOFLGFRSFLKILTCMOGHEETIRROGHLEVDPDMEAAIAIQM 540
Qy 541 QLNKILLMFOEWACDEBELLVAYKECHKAVNRCSTPSISSKTVVQSGHSLKTSYRV 600
Db 541 QLNKILLMFOEWACDEBELLVAYKECHKAVNRCSTPNFSSKTVVQSGHSLKTSYRV 600
Qy 601 SEDLSVHLPLSRTLAGLHVRLSRLGAVSRHLEFVSFQVEVLVEYPLRCLVLVAQV 660
Db 601 SEDLSVHLPLSRTLAGLHVRLSRLGAVSRHLEFVSFQVEVLVEYPLRCLVLVAQV 660
Qy 661 AEMWRNGLSLISQVFPYQDVKREMYDKDILMLQIGASLMDPNKFLILVLORVELAE 720
Db 661 AEMWRNGLSLISQVFPYQDVKREMYDKDILMLQIGASLMDPNKFLILVLORVELAE 720
Qy 721 ENKTIKTDODLKOYNTLIEEMLOVLIYIGERYVPGVGNVTKBEVTMREIHLCLIEP 780
Db 721 ENKTIKTDODLKOYNTLIEEMLOVLIYIGERYVPGVGNVTKBEVTMREIHLCLIEP 780
Qy 781 MPHSAIARLNENNETGLENVINKVATFKPGVSGHGVYELKDBESLKFNNFYYSK 840
Db 781 MPHSAIARLNENNETGLENVINKVATFKPGVSGHGVYELKDBESLKFNNFYYSK 840
Qy 841 TQHSKAEHMQKRRQKQENKORALPPPPPPPCPAPSKVINLNCIDIMWYILTVTERAID 900
Db 841 TQHSKAEHMQKRRQKQENKORALPPPPPPPCPAPSKVINLNCIDIMWYILTVTERAID 900
Qy 901 TDSNLTWTEGLOWAPHILALGALLEKQOLKAPBEERTDFYHKASRLGSSAM--NIQM 957
Db 901 TDSNLTWTEGLOWAPHILALGALLEKQOLKAPBEERTDFYHKASRLGSSAMNAQIQM 960
Qy 958 LLEKLGIPQLEGQKDMITWILQMPDVTVRLEKSLIVATTSGESIKNDITHDKKA 1017
Db 961 LLEKLGIPQLEGQKDMITWILQMPDVTVRLEKSLIVATTSGLEIKSEBITDKKA 1020
Qy 1018 ERKRAEAAHLRQKIMQASALQKRFIETHKLMYDNTSEMPGKEDSIMEERSTAVSDY 1077
Db 1021 ERKRAEAAHLRQKIMQASALQKRFIETHKLMYDNTSEVTGKEDSIMEERSTAVSEA 1080
Qy 1078 SRIALGPKRGPSTVEKEVLTCILCQEEQVKIENNAWLSACVQKSTALTQHRGKPIELS 1137
Db 1081 SRIALGPKRGPSTVEKEVLTCILCQEEQVKIENNAWLSACVQKSTALTQHRGKPDHL 1140
Qy 1138 GRALDPLFMDPDLAYGTYTSGCGHVMHVCWKYFEAVQLSSQORIHVDLPLESGEYLC 1197
Db 1141 GETLDPLFMDPDLAGHTYTGCGHVMHVCWKYFEAVQLSSQORIHVDLPLESGEYLC 1200
Qy 1198 PLCKSLCNTVPIPILOPKINSNADALAQLLTLARWITQTLARI SGYNIRHAKGNP- 1256
Db 1201 PLCKSLCNTVPIPILOPKINSNADALAQLLTLARWITQTLARI SGYNIRHAKGEAPA 1260
Qy 1257 IPIFFNQMGDSFLRPHSILSGVRESSIKYSNSIKEMVILFATTIRIGLKVPPDERDP 1316
Db 1261 VPVLFNQMGDSFTFRPHSILSGVQSSVKYSNSIKEMVILFATTIRIGLKVPPBELDR 1320
Qy 1317 VPMJLWSTCAFTIQAIENTLIDGEGKPLFGALQNRQHNGKALMQPVAQVARIITCPOVLIQ 1376
Db 1321 VPMJLWSTCAFTIQAIENTLIDGEGKPLFGALQNRQHNGKALMQPVAQVARIITCPOVLIH 1380
Qy 1377 HLVRLLSVLPLNKSDETPCLLSIDLPHVLVGNVLA PPSLYWDDPDVQLPSSVSSYNHL 1436
Db 1381 HLARLLSVLPLNKSDETPCLLSIDLPHVLVGNVLA PPSLYWDDPDVQLPSSVSSYNHL 1440
Qy 1437 YLFHLITMAHMLQILLTVDVTDV---GLPLAQVQDSEAHASFPFABISQVTSIGGCDI 1492

Db 1441 YLFHLITMAHMLQILLTVDVTDVTDLSPPPLAEGEEDSEARCASAFVEVSQHTDGLTGCGA 1500
Qy 1493 PGWYLVSLKNGITPYLRCAALPHYLLIGVTPPBELHTNSAEGVYSA LCSYLSLPTNLPL 1552
Db 1501 PGWYLVSLKNGITPYLRCAALPHYLLIGVTPPBELFANSABGFSALCSYLSLPTNLPL 1560
Qy 1553 LFOEYWDVTRPQLQWACADPALLNCLKQKNTVVRYPKRNLSLIELPDDYSCLLNQASHFR 1612
Db 1561 LFOEYWDVTRPQLQWACADPALLNCLKQKNTVVRYPKRNLSLIELPDDYSCLLNQASHFR 1620
Qy 1613 CPRSADDERKHPVLCLFCGAILCSQNICCOBIVNGEYVACIFHALHCGAGVCIPLKIRE 1672
Db 1621 CPRSADDERKHPVLCLFCGAILCSQNICCOBIVNGEYVACIFHALHCGAGVCIPLKIRE 1680
Qy 1673 CRVVLVEKARGCAYPAPYLDYGETDPLKRGNPLHLRSRERYKHLVMOQHCHIEIA 1732
Db 1681 CRVVLVEKARGCAYPAPYLDYGETDPLKRGNPLHLRSRERYKHLVMOQHCHIEIA 1740
Qy 1733 RSQETNOMLFGFNWQLL 1749
Db 1741 RSQETNOMLFGFNWQLL 1757
RESULT 4
US-09-724-126A-4
; Sequence 4, Application US/09724126A
; Patent No. 6706505
; GENERAL INFORMATION:
; APPLICANT: Han, Hui-Quan
; APPLICANT: Kwak, Keith
; TITLE OF INVENTION: Human E3 Alpha Ubiquitin Ligase Family
; FILE REFERENCE: 01017/35966A
; CURRENT APPLICATION NUMBER: US/09/724,126A
; CURRENT FILING DATE: 2000-11-28
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 1755
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-126A-4
Query Match 46.9%; Score 4327; DB 4; Length 1755;
Best Local Similarity 47.2%; Pred.No. 0;
Matches 841; Conservative 339; Mismatches 543; Indels 58; Gaps 27;
Qy 1 MADE---BAGGTER--MEISAEIPQTPQRLASWMDQVDVYTAFLHLAQLVPEYFAEM 55
Db 1 MASELEPEVQALDRSLLECSAE-----EIAKWLQATDLTREVYQHLAHVYPKIYCRGP 54
Qy 56 DPDLKQESVQMSIFTEPLEWYLFGEEDPDI CLEKLKHSQ-AFQLCGRVFKSGETTYSCRD 114
Db 55 NPFPQKEDMLAQHVLGLPMWYLCGEDPAFGPKLEQANKPSHLCGRVFKGVEPTYSCRD 114
Qy 115 CAIDPTCVLCHMDCFODSVHKHRYKMYHTSTGGFCDCGDTAWKTGPPCVNHEPBRAG-T 173
Db 115 CAVDPTCVLCHMDCFODSVHKHRYKMYHTSTGGFCDCGDTAWKTGPPCVNHEPBRAG-T 174
Qy 174 IKENSRCPLNEVIVQARKIPPSVIVYVEMTWEEKELEPPQLQIRKKNRYCYCVLND 233
Db 175 BEEDPLVHLSVDVIARTNIFATFRYAVAILTWKESELPALEWVEKSDTYICMLFND 234
Qy 234 EHSYDHYVYLSQALDCELABQAHLHTTAIDKGBRRAVAGAYAAQCBAKEDIKHSENV 293
Db 235 EVHTYEQVIYTLQKAVNCTQKEAIGFATTVDRGSRVRYGDFQCEQAKSVIVRNTSRQ 294
Qy 294 SOHPLHVEVLHSEIMAHQKFAIRLGLSWNKNIMSYSSDRFQICQACLREEDSENPLIS 353
Db 295 TK-PLKVQVHMSSIVAHQNFGLKLSLGLSIIIGSDGLRILLCVQLGEGPGDGENSLVD 353
Qy 354 RLMLWDKLYKGARKILHELIFSSPFMEYKGLFAMEFVKYKQLOKEYISDDHDSIS 413

354 RLMLSDSKLWKGARSYHQLFMSLLMDLKYKLPFAVRFAKNYQQQORFDMEDDHERAVS 413
414 ITALSVMFTPTVPLARHLEEQNVISVITETLEVLPEYLDRNNKFNFOGYS---QDKLG 470
414 VTALSVOFTPTAPTARMLTEENLMSIIKTMDHL-RHRDAQGRFOFERYTALQAFKER 472
471 RVYAVICDLKYILISKGTITWTERLMQPLEGFRSFLKILTCMQGMEIERBROVQOHIETDP 530
473 RVQSLDLIDKYILISKGTITWTERLMQPLEGFRSFLKILTCMQGMEIERBROVQOHIETDP 532
531 DNEAALAIQWOLKNVILQEWACDEBELLVAYKECHKAVMRCSTSFSSKTVVQS-C 589
533 EWEAAATLQWLTWLSIMQWDCASDEKVLIEAYKCLAVLMQCHGYTDGEPITLSC 592
590 GHSLETKSYRVSBDLVSIIHLPLSRTLAGLHVLRLSGLAVSRLEHFPBDFQVEVLVEYP 649
593 GHSVETIRYCVSQEKVSIHLPLVSRLLAGLHVLRLSGLAVSRLEHFPBDFQVEVLVEYP 652
650 LRCLVLVAQVAVMWRNGLSLISQVFTYQDVKCREMYDKDIIMLOIGASLMDPKPFL 709
653 LRCLVLCAQVHAGMWRNGLSLISQVFTYQDVKCREMYDKDIIMLOIGASLMDPKPFL 712
710 LVLORVELABAFN---KTISTK---DODLIKOYNTLIEEMQLVLIYIGERVVPGVNV 762
713 IMLSRFELQISTPDYGRFSEITHKQVQNNNTLIEEMQLVLIYIGERVVPGVNV 772
763 TKEEVTMBRIIHLCEPMHSAIAKNLPENNENETGLENVINKVATFKKPGVSGHYVE 822
773 NATDEIKREIHLQSLIKMAHSELVKSLEPENKETGMSVLEAVAFKKPGLTGRGMYE 832
823 LKDESILKDFNMFTYHYSKQHSKAEHMQKRRKQENKDALPPPPPEPCAPSKVINL 882
833 LKPECAKBNFLYTHFSRAQSKAEBAQRKLRQNRREDTALPPVLPFPFCLFASLVNLT 892
883 NCDIMMYILRTFERAIDTDSNLWTEGMLQMAFHIALGLLEKQOQKAPREE-VTPDP 941
893 QSDVMLCINGTLQWAVEHNGYAWSESMLQVHLHGLMALQOEKQOHLNTEBEHVTPTF 952
942 YHKASRLGSSAMN---IOMLEKLGIPOLGOKOMITWILQWDFTVKRLREKS-CLIVA 997
953 TQKISKPGEPKNSPISLAWLETQWAPYLEVHKOMIRWILKTFNVAKORSSPSPVA 1012
998 TTSSESINKDEITHDKBAERKAKABAARLHRQKIMQMSALQKNFIETHKLMYDNTSE 1057
1013 ETEGTTM---EESSRDKDAERKAKABAARLHRQKIMQMSALQKNFIETHKLMYDNTSE 1069
1058 MPKEDSIMEESTPAVSDYSRIALGPKGSPVTEKEVLTCTILCOEQQVKLENNAMVLS 1117
1070 LQASTSAVLH---SPVASDMTLTALGPTQTPQEQRFVTCILCOEQQVKLENNAMVLS 1127
1118 ACQVKSTALTQHRGKPIELSGBALDPLFMDPLAYGTYTGSCHVNVHVCWKQYFEAVOL 1177
1128 AFQVQSTVLKRSKRFQI-DPEKYDPLFMDPLAYGTYTGSCHVNVHVCWKQYFEAVOL 1186
1178 SQQQ-----RIHVDLPDLSGEYLCPLCKSLCNTVPIIPILOPKINSENADALAQLLPL 1232
1187 KEORRQORLRLHYS-YDVENGEBFLCPLCELSNTVPIIPILOPKINSENADALAQLLPL 1243
1233 ARWQITVLARIISGYNIRHAKGENPIPIFNQMGSDSTLEPHSLISFGVSESSIKYNSIKE 1292
1244 TQWIRTISQOIKALQFLRKEESTPNNASTKNSENVDELQPLGPRFDPFAPKIPYSESIKE 1303
1293 WYLPATTYIRIGLKVPPDERPRVPMLTWSTCAFTIOAIENLGLDGBGKPLFGALQNRQH 1352
1304 MTTTGTATYKVLGKVPNEEDPRVPMCMGSCAYTIQSIERILSDSDKPLFGPLPCRLD 1363
1353 NGLKALMQFAQRITCPOVLIOGLHVRLLSVLPIKSEDTFCLISIDLFHVLGAVLA 1412
1364 DCLRSITRFAAAHWTVASVSVQGHFKLPLASIVPNDSEELPCILIDMFHLLGLVLA 1423
1413 FPLSLYWDPPDVLQPSVSSVSNHLYLFLHITMAHMLQIILL---TVDTLPLAQVQDSESE 1469

1424 FPALQCOD-----PSGISLGTGDLHFLHVLVMAHIIQIILLTSTCTENGMD--QENPPCEE 1476
1470 AHSASFFABISQYTSIGSGDIP-GWYLVMSLKNIGITPYLRCAALFFHYLLGVTPPEEL 1528
1477 EBAVALALYKTLQYOT-GSALKIPESGHMLWSVRAGIMPLFKCSALFFHYLNGVPSPPDI 1535
1529 HTNSAEGEYSALCSYLSLPTNLPLFOEYWDVTRPQLQWCADPALLNCLGKQNTVVRYP 1588
1536 QV-PGTSHPHEHCSYLSLPTNLPLFOEYWDVTRPQLQWCADPALLNCLGKQNTVVRYP 1594
1589 RKRNSLIELDDYSCLLNOASHFRCPRSADDSKRPVLCFLCGAILCSQNICQETVNGE 1648
1595 RESNKLINLPEDYSSILNOASHFRCPRSADDSKRPVLCFLCGAILCSQNICQETVNGE 1654
1649 EYVACIFHAIHCGAGVCIPLKIRRCRVLVVEGKARGCAYPAPYLDYBGTDPGLKRGNPL 1708
1655 DVGACTAHTYSCSGVGIFLRVRECVLPLAGTKGCFSPYLDYBGTDPGLKRGNPL 1714
1709 HLSRERYKHLVWQOHCIIIEIARSQETNOMLFGFNWOLL 1749
1715 HLCKERFKIKQKLWHOHSVTEIGHAQEANOQLVGDIMQHL 1755

RESULT 5
US-09-724-126A-6
; Sequence 6, Application US/09724126A
; Patent No. 6706505
; GENERAL INFORMATION:
; APPLICANT: Han, Hui-Quan
; APPLICANT: Kwak, Keith
; TITLE OF INVENTION: Human E3 Alpha Ubiquitin Ligase Family
; FILE REFERENCE: 0101735966A
; CURRENT APPLICATION NUMBER: US/09724,126A
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 60/187,211
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 6
; LENGTH: 1755
; TYPE: PRT
; ORGANISM: Mouse
US-09-724-126A-6

Query Match 46.4%; Score 4278; DB 4; Length 1755;
Best Local Similarity 46.5%; Pred. No. 0;
Matches 835; Conservative 341; Mismatches 534; Indels 84; Gaps 28;

QY 1 MADE---BAGGTER--MEISAEPLPOTPORLASWQDQVDFYTAFLHLLAQLVPBIYFAEM 55
DB 1 MASEMEPEVQAIDRSLLCSAE-----EIAGRWLQATDLNREVYQHLAHCVPKIYCRGP 54
QY 56 DPDLKQESVQMSIPTPLEWYLFCEDPDIDCLEKLKHSQ-AFOLCGRVPKSTETYSCHD 114
DB 55 NPFPOKEDTLQAHLILGPMWYICABDPALGPFKLEQANKPSHLGCRVPKVGEPYSCND 114
QY 115 CAIDPTCVLCMDCFQDSVHKHRYKMTSTGGFCDCDGTBAWTKGPPFCVNHPEGRAGTI 174
DB 115 CAVDPTCVLMCEFLGSIHRDHYRMTTSGGGFCDCDGTBAWTKGPPFCVNHPEGRAGTI 174
QY 175 -KENSRCPLNEVIVQARKIPPSVTKYVVMETIWEKEKELPPLQIRKERNRYCVLNFND 233
DB 175 BEEDPLVHLSDEVARTNYTIPALPRYAVDILTWKESSELPLEVASKSDTYTCWLFND 234
QY 234 BHSYDHYVLSIQRALDCELAQLHTTAIDKEGRVAKAGAYACQAKEDIKSHSENV 293
DB 235 EVHTYEVIYTLQKAVNCTQKBAIGFATTVDRDGRPRVRYGDFQCDQAKTVIVRNTSRQ 294
QY 294 SOHPLRVHLSHSEIMAHQKVALRLGSMNKMYSYSSDFROIQACLRPEEPSENPLIS 353
DB 295 TK-PLKVVQWSSVAHQNFGLKALSLVGSVIGSDGLRRLILQVGLQEGPDGESSLVLD 353
QY 354 RLMLWDAKLYKGARKILHELIFSSPFMEYKCLFAMBFVKYKQKQKEYISDDHRSIS 413

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 45019

LENGTH: 334

TYPE: PR

ORGANISM: Drosophila melanogaster

US-09-270-767-45019

Query Match 4.9%; Score 450.5; DB 4; Length 334;

Best Local Similarity 35.3%; Pred. No. 1.1e-31;

Matches 108; Conservative 48; Mismatches 103; Indels 47; Gaps 10;

QY 1463 VORDSE--AHS-----ASSFPAISQYTSIGCDIPGWLWVSLKNGITPYLRCA 1513

DB 51 VESEENQHGQTVIPCESHILALLEY-----VQMSFLRCS 93

QY 1514 LPHYLLGVTPPELHTNSAGBSYALCSYLSLPTNLFLFQBYMDTV-RPLQRCWADP 1572

DB 94 LFRFLTDVDFPTDPTDQPD-RFDLMQYGLDPLMGVYFD--METVYATMHSFASHP 150

QY 1573 ALLNCLKOKNTVYRPRKNSL-IE-----LPDYSCLLNOASHFRCPRSADDR 1620

DB 151 HIDREVEQR-----CQDPAARRSLQVPCPLRPLKVLKLCDDFSLINSVSDIFCPNNREE 206

QY 1621 RHPVLCLFCGAILCSNICCOBIVNGEEVGACIFHALHCGAGVCIFLTKRECRVVLVSG 1680

DB 207 MKTPTWCLLGLLCCQSQSCCQPELKVSVGACTHEAHAGAEVGLRIRDCQVYL-G 265

QY 1681 KARGCAYPAPYLDEYGETDPLKRGNPLHLRERYRKLHLVMOQHCHIEBIARSQETNOM 1740

DB 266 RGKGCVPVPPYLDYGETDGLRGNPLRLSQAYRYKIYQLWLHGLHGLAEIARLNDNAV 325

QY 1741 LFGFNW 1746

DB 326 AAAAQW 331

RESULT 8

US-09-248-796A-19402

Sequence 19402, Application US/09248796A

Patent No. 6747137

GENERAL INFORMATION:

APPLICANT: Keith Weinstock et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN

FILE REFERENCE: 107196.132

CURRENT APPLICATION NUMBER: US/09/248,796A

CURRENT FILING DATE: 1999-02-12

PRIOR FILING DATE: 1998-02-13

PRIOR APPLICATION NUMBER: US 60/096,409

PRIOR FILING DATE: 1998-08-13

NUMBER OF SEQ ID NOS: 28208

SEQ ID NO 19402

LENGTH: 817

TYPE: PR

ORGANISM: Candida albicans

US-09-248-796A-19402

Query Match

Best Local Similarity 22.2%; Score 317.5; DB 4; Length 817;

Matches 205; Conservative 125; Mismatches 327; Indels 268; Gaps 42;

QY 961 KLKGIPLQEGQDMITWILQMPDTVKRLKSKLIVATTSGSEIKNDREITHDKKAERK 1020

DB 25 KRSIFKINQK-----YSPFTPLKARSEFDISKVDLSDDANSEDLGKKRIAKER 79

QY 1021 RKAAARLHRQKIMQMSALQKRNFIETHKLMYDNTSEMPGKEDSIMEBSTPAVSDYSR 1080

DB 80 QAKLMAKFKQ-----QSLFLKKNQPET-----I 103

QY 1081 ALGPRKGPVSVEKEVLT-----CILQEQEVEKIENAMVLSACVOKSTALTOH 1129

DB 104 ACG-----SDTEMBELDDHAAWKPPPEHHCMLCQNAE-----DAGPFGITAYISK---CBF 152

RESULT 9

US-09-248-796A-19403

Sequence 19403, Application US/09248796A

Patent No. 6747137

GENERAL INFORMATION:

APPLICANT: Keith Weinstock et al

TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA

FILE REFERENCE: 107196.132

CURRENT APPLICATION NUMBER: US/09/248,796A

CURRENT FILING DATE: 1999-02-12

PRIOR APPLICATION NUMBER: US 60/074,725

PRIOR FILING DATE: 1998-02-13

PRIOR APPLICATION NUMBER: US 60/096,409

PRIOR FILING DATE: 1998-08-13

NUMBER OF SEQ ID NOS: 28208

SEQ ID NO 19403

LENGTH: 631

TYPE: PR

; ORGANISM: Candida albicans
US-09-248-796A-19403

Query Match 3.0%; Score 274.5; DB 4; Length 631;
Best Local Similarity 22.2%; Pred. No. 2.4e-15;
Matches 154; Conservative 107; Mismatches 232; Indels 201; Gaps 34;
QY 598 YVSEDLVSIHLPLSTAGLHVRLSRLGAVSRL-----HEVVSFEDFOVEVLVEYPLRC 652
DB 47 FOVSKEPVAIVNPLN-SLLSYLQFQNDNFKHLQQNOSLVNFTDIS-----LRS 97
QY 653 LVLVAQVAMWRNGLSIIQVFFYQDVKCEEMVDKDIIMLQICASLMDPKFKLLAVL 712
DB 98 IVLGSQVAGFWRNGSSVSROATAYFDSLMTAYAFMRDPHLNQAILPDEQKVLNPL 157
QY 713 QRYELAE-----AFNKTI-----STKDDILIKQNTLIEMLQVLIYIVGERVPGVG 760
DB 158 ERWELRSWPKNEVDFDKTIYERFPIVERFTAFAYNLFVDRSMFI----- 203
QY 761 NUTKEVMTREIHH-----LLCIEPMPHSAIAKNLLENNETGLENNVINKVATPKP-CV 815
DB 204 NETPEDATRLKLNHAIGYALCEAMAFSELHQYIDTNIASSKFDVLVEYADYQPPSAL 263
QY 816 SGHGVVELKO---ESLKDPNMYF---YHYSKTQHSKAERHMQKRRKQENKDEALPPP- 867
DB 264 TDSGLYRLKESTYEKLDPLNLLVDCKFQIISEMWIKNSKQKRR-----KSENLIIVTPV 318
QY 868 -----PPFCPAPSKVINLNCDDIMMYIIRTVPERAIDTDSNLWTEGMLQMAFHILA 919
DB 319 IEKAGNDFVDENIGNPAKSLPFVK---LIYKLMQV---SIDTSDETYLPHLLHLV-HAIM 371
QY 920 LG--LLEEKQOQKAPEE-----EVPDFY--HKASRLGSSAMNIQMLLEKLG 964
DB 372 LDDEMIHGKEYLNKHFPVDIPITDLLTILESTMSKYVCOKADYL----- 415
QY 965 IPQLEGO-KDMITWILQMF--DTVKRLREKSLIVATTSGSISIKNDIETHDKKAERKR 1021
DB 416 VQLGVGKDRRIIASLVDCGEDIYQYKRR-----KNSLSPESAER--KKR 459
QY 1022 KAAARLHRQKIMAOQSALQKNPIETHKLMYDNTSEMPKESIMEEESTPAVDYSRIA 1081
DB 460 KAEK---RKNPIPKKSKQREKFLNQ---BFQLDIPQDSNIPMEGDS----- 503
QY 1082 LGPKRGPSTVEKVLTCILCOEEQEVKINNAMVLSACVQKSTALTQHRGKPIELSGEAL 1141
DB 504 -----RKLRCTVACGELE--SFEKPLAIMAA-----STKAPVFWKVPVQ-TGEVV 545
QY 1142 DPLF-----MD-PDLAYGT-----YTGSCGHVHVAVCWQKYFRA 1174
DB 546 SNAFTWKNILLDYKDTYGVGYDPVSRSLSDSNRPESYVLSTCGHSIHHSCLNR---- 601
QY 1175 VQLSSQORIHVLDLDESGEYLCPLCKSLCNTVI 1208
DB 602 -----RI-----IGSAQYSCPLCHNLHDMII 622

RESULT 10

US-09-270-767-32482
; Sequence 32482, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 32482
; LENGTH: 280
; TYPE: PRT
; ORGANISM: *Drosophila melanogaster*
US-09-270-767-32482

Query Match 2.1%; Score 197; DB 4; Length 280;
Best Local Similarity 26.1%; Pred. No. 5.8e-09;
Matches 81; Conservative 38; Mismatches 127; Indels 64; Gaps 11;
QY 1054 NTSEMPGKEDSIMBEESTPAVDYSRIA-LGPKRGPSTVEKVLTCILCOEEQEVKIENN 1112
DB 1 SASTGPDWEDI PPBEEQGAVALSKVACIAGPDRKFYHGTDDTPKCILCFENCAISRGR 60
QY 1113 AMVLSACVQKSTALTQHRGKPIELSGEALDPLFMDPDL---AYGTYTSCGHVHVAVCWQ 1169
DB 61 OLVSASFQTSRV-----IFTPNLRNSQSALHISCCGHVHYSCLW 102
QY 1170 KYF---BAVOLSSQORIHVDLFDLESBYLCPCKSLCNTVPIPILOPQ----- 1216
DB 103 EYFTNEEFKELRPHRNRRAALAAQANVEFQCPYCNLTSLNAIIPVTETLPASAPSPNES 162
QY 1217 -----KINSENADALAQL-----LTLARWIQTVLARISGYNIRHAKGENPIPIFFNQG 1264
DB 163 YLPDLSFVEINSTLAIELGNVKDHETLTPSVSNIL-ELSGVVGGLAQFERSVOLIKNPP 221
QY 1265 MGDSTLFHSHLSFGVSESSIKYSNIKEMVILPATTIYRIGLK-VPPDERDPRVPMLTWS 1323
DB 222 -----RLHADYIEGIE-----FLKKALLNTMKIOQSHLKHDPALIESIEMVPII-WD 266
QY 1324 TCAPTIOAIE 1333
DB 267 SCSYTLQALE 276

RESULT 11

US-09-949-016-6507
; Sequence 6507, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6507
; LENGTH: 3259
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6507

Query Match 2.1%; Score 190.5; DB 4; Length 3259;
Best Local Similarity 18.0%; Pred. No. 1.4e-06;
Matches 221; Conservative 207; Mismatches 424; Indels 377; Gaps 49;
QY 164 VNHEPFRAGTTIKENSRCPLNEEVIQARKIPPSPVIKYVEMTIWBEKEKLPPELO--IRE 221
DB 12 VLHELSDGDDTDDQNRAFLDPELHQESDMEFNNTTQEDVQERLAVAE-QLVVVELKDIIRQ 70
QY 222 KNERYYCVLFNDEHSHYDHVYISLQALDCELAQAQLHT-----TAIDK--EGERRAVKAGA 275
DB 71 KDVLQ-----QQKDEALQBERKAADNKIKLKLHAKAKLTSLNKYIEMKAGGGTV 121
QY 276 YAAQOEAKEDIKSHSENVSOHPLHVEVLHSEIMAKKPAKRLGSGWNWTKIMSYSSDFRQIF 335
DB 122 LPTFQSESEQSKHDKSSTSEMEETEKIKHLOEKEELISTL----- 163
QY 336 COACLRBEPDSNPCLISRLMLWDALYKGARKILHELIFSSFFFMEMEYKCLFAMEFVKY 395

Db 164 -QAQL-TQAQEAQPAOSS-----TEME-----PFVMM 188
Qy 396 YKQLOKEYISDDHRSISITALSVMQFTVPTLARHLIEBON-----VISVIT 442
Db 189 KQQLQEK-----EEFISTLAQLSQTAQEAQAQVVRREKDAFETQVRLHEDELQVLT 242
Qy 443 ETLLVLPEYLDNRNKNFQGSQDKLGRVYVICDKYLILIS-----KPTIWTERRLM-- 496
Db 243 QADVETEMQOKVLQRKLEHEESLVGR--AQVDDLOQELTAASORNOILSQQLOQME 300
Qy 497 -----QFLEGRSFLKILTCMGMEERROVGHIIEVDPDWEAAIAIQMOLKNILLMP 549
Db 301 AEHTLRNTVETEREESKILLEKWELEVAERKLSFH-----NLQEMGHILLEQF 349
Qy 550 QEWACDEBL--LLVAYKECHKAVMRCSTFSI--SSSKT--VVOGCHSLETYSYRVSDEL 604
Db 350 EQAGQAQAELEGRYSALQKHAEMBEKTSILSLQKTGOELQSDACDALKQNSKLLQDK 409
Qy 605 VSHLPLSLTLAGLHVLSRLGA-----HSP-----VSPEDFOVEVLVEYPLRC 652
Db 410 NEQAVQSAQTIQOELQLOQKSKBSIOFLNRLPQLOQHETASQTSFPDVYNEGTAVENTEEN 469
Qy 653 LVLVAQVVAEMWRNGLSLISOVFYQDVKCREEMDYKDIIIMLOI-----GAS 700
Db 470 IASLQKRVLENEKCALILSSI-ELEBELKAENEKLSQITLLEAQRNTRGEADREVSIS 528
Qy 701 LMD-PNK-----FLLLVLQRYELARA 720
Db 529 IVDIANKRSSAESQDVLENTFSQKHKELSVLLEMEKAEQEEIAFLKQLQKREABEA 588
Qy 721 FNKTISTKQDOLIKY--NTLIEMLQVLIYVGERY----- 755
Db 589 DHEVL---DQEMKQMEGIAPIKMKVFLDTGDPFLMPNEESSLPAVEKEQASTHQ 645
Qy 756 -----VPGVGNVTKERTVREIILHLCIEPMPHSA- 785
Db 646 STSBEISLNDAGVELKSTQDQDLSLSAVPDIGQCHQDELB-RLKSQILELELNFHKAQ 704
Qy 786 --IAKNLPENNETGLENVINKVATPKPGVSGHGV-----ELKDESL---KDFNMYP- 835
Db 705 EYERKULDEKAKEISNLQLIBE---FKGNADNNSAFTALSEERDQLLSQVKELSMVTE 761
Qy 836 --YHYSKTOHSAEAMQKGRQENK---DEALPPPPPPFCFAPSKVINLNCIDIMVY 890
Db 762 LRAQVKQLEMLAEARORRLDYESTAHNDLLTE-----QIHSLSTEAKSKDVKIEV 814
Qy 891 LRTVE-----RAIDTDSNLWTGMLQMAFH 916
Db 815 LQNELDDVOLQFSQBSTLRSIQLQNKSEVLEGAERVRHISKVBELSQALSQKELE 874
Qy 917 ILALG--LLEBK---QOLQKAPEE-----EVTDFYHKASRLG---SSAMNIQMLLEK 961
Db 875 ITKMDQLLEKRDVETLQATIEEKDQVTEISFSPMTKRWQLNEBKSLGVEIKLKEQ 934
Qy 962 LKGI-----PQLEGQDMITWILQMDPTVVRUREKSLVATTSSGESIKNDIEITHD 1013
Db 935 LNLLSRAEBAKKEQVEEDNEVSGLKQNYDEM-----SPAQISKEBELQHE 980
Qy 1014 -----KEKAERKKAARLHQIKMAQMSALQKPIETHKLWYDNTSEMPKESIMEE 1068
Db 981 FOLLKKEQORRKLQALINKEKLLQVRSLRE---ELANLKDSEKKEIPLSETERGEV 1037
Qy 1069 ESTPAVSDYSRIALGPKRGPSTVEKVLTCILQCE-----EOEVKIENNAWLS- 1117
Db 1038 EEDKENKEVSE-----KCVTSK-----CQIEIYVKQITISEKEVELQHIRKDLER 1082
Qy 1118 --ACVOKSTALTOHRKPKIELSGREALDPL 1144
Db 1083 KLAABEQFQALVKMONTQDNTQIDLL 1111

RESULT 12

US-09-949-016-6468

; Sequence 6468, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6468
; LENGTH: 1979
; TYPE: PRT
; ORGANISM: Human
; US-09-949-016-6468

Query Match 2.0%; Score 184; DB 4; Length 1979;

Best Local Similarity 18.7%; Pred. No. 2.3e-06;

Matches 233; Conservative 195; Mismatches 453; Indels 366; Gaps 53;

Qy 209 BEKLEPPELQREKNERYCVLFDNHHSHSYDHYVLSQRLDCELAEAQLHTTAIDKEGR 268
Db 79 EKHASHIQIKQOSTSTANQL-QQKEVEISH-LKARIALQDQLLKQ-----SNA 128
Qy 269 RAVKAGAYAAQCAKEDIKSHSNYSOHP--LHVEVL-HSEIMAHQKPAIRIGSMWNKIM 325
Db 129 QSVPSGAGVPATTAS---SSPAYGISHPSPAFHDDMDFGDIISQOEINRLSNEVSRLE 185
Qy 326 SVSSDFRQIPQACLRBPDPSENPCLSRLMLWDAKLYKGARKILHELIFSPFFMEMEYK 385
Db 186 SEVGHWRH-AQTSKAQGTDSNDQSEICKL----- 214
Qy 386 KLAFMEFVKYKQLOKBYISDDHRSISITALSVMQFTVPTLARHLIEBONVISVITEL 445
Db 215 -----QNIILKELKNRSQEI-DDHQEMSVLQNAHQKQTEISRRHRELSDY-----EER 264
Qy 446 LEVLPEYLDNRNKNFQGSQDKLGRVYVICDKYLILSKPTIWTERRLMQFLGFRSF 505
Db 265 IEELNLLQGG---SGVIEDLSKIYEMOKTIQVLQIEK--VESTKOMEQLEDKIKDI 318
Qy 506 LKILTCMGMEERROVGHIIEVDPDWEAAIAIQMOLKNILLMPQEWACDEBELLVAYK 565
Db 319 NKCLSSAENDRDLRREQEQLNVEK-----RQIM-----E 348
Qy 566 ECHKAVMRC---TSPISSTKTVQSCGSHLETYSYRVSDELVSITH--LPLSRTLGLHV 620
Db 349 ECENLKLSCSKLQPSAVKQSDTM-----TEKERILASASVEVFRLOQALSDARN 399
Qy 621 RLSRIGAVSRILHFPESFEDFQVFLVBYRLCLVLVAQVVAEMWRNGLSLISQVIFYQ- 679
Db 400 EIMRLSSLNQNSLAE-DNLKLMKRIEVLKEKSLLSQKEEL-----QMSLLKLANVEV 454
Qy 680 -----DVKCREEMYDKDIIIMLOIGASIMDPNKL-----LLVLQRYEL-----ABAFNKT 724
Db 455 IKSTATRDISDSELHD---LRLANLEAKEQELNQSISEKETLIAEIEBELDRONQEAATKHM 511
Qy 725 ISTDQDOLIKQYN---TLIEBMLQVLIYVGERYVPGVGNVTKERTVREIILHLCIEBPM 781
Db 512 ILIKDQ-LSKQONEGDSIISKQLQDL-----NDEKKRV-----HQLEDDKM 551
Qy 782 PHSAIAKVLPENNETGLENVINKVATPKPGVSGHGVYELKDESLKDFNNFYHYSKT 841
Db 552 D---ITKELDVQKEKLIQSEVALNDLHLTKQK-----LEDKVENLVD-----QLNKS 595
Qy 842 QHS-----KAEMQKGRKQENKDEALPPPPPPFCFAPSKVINLNCIDIMMYILRTVFER 897

```
Db 596 QESNVSIQENLELKEHIRONEE-----LSRIRNEL-----MQ 629
Qy 898 AIDTDSNLWTEGMLQWAFHILALGALLEBEKQOQKAPK--EVTDFYTHKASRLGSSAMNI 955
Db 630 SLNODSNSNFKDTLLKEREAEVRNLKQNLSELEQLNENLKVAFDVKMENKLVACEDV 689
Qy 956 QMLLEK-LKGIPOLEGQKDMITWILQMPD-----TVKRLREKSCLIIVATTSGSESI 1005
Db 690 RHQLEBCLAGNQLSLEKNTIVETLKEBGEIEABLCKWAKLLBEANKYKTEIELSNA 749
Qy 1006 KN-----DBITH-----DKEKAERKRKAE-----AARLHQKIMAQM 1037
Db 750 RNLTNSALQLEHEHLIKLQKKDMEIAELKKNIEQMDTDHKETKDVLSSELEQKLTOL 809
Qy 1038 SALQKNPIETHKLMYDNTSEMPGK-----EDSIMEEES 1070
Db 810 INKKEIFIEKLRKERSKQLELDKYSQLRKNNEILRQTIEBKDRSLGSMKEENHLOEEL 869
Qy 1071 TPAVSDYSRIA--LGPKRGPSTVE--KEVLTCILQOEBOEVKIENNAMVLSACVOKSTAL 1126
Db 870 ERLREEQSTAPVADPKTLDSTVELASEVSLNTIKHELEBEIKHKQKIIEDQNSKQML 929
Qy 1127 TQHRGKPIELSGEALDPLFMDPDLAYGTYTSGCGHMHAVCQKYPEAVQLSSQQRHVD 1186
Db 930 LQ-----SLQEQKEMDEF-----RY-----QHEQMATHQ 956
Qy 1187 LPDLSGEVLCPLCKSLCNTVPI-----IPLQPKINSENADALAQLLTLARWITVLARI 1243
Db 957 LP-LEKDEBI-----KSLQKTIIEQIKTQLHEERQDIQDTSNDFQEB-----TKV 999
Qy 1244 SGYNIRHAKGENPIPIFFNQGMGDSITLPHSILSGVSESIKYNSIKEMWILFATTIYR 1303
Db 1000 QSLNIENG-----EKHDLKSKATERLVK-----1023
Qy 1304 IGLKVPDPDRPRVPMLTWSTCAFTIQAIENTLGDGKELFGALQNR 1350
Db 1024 -GIK-----BRELEIKLNEKNISLTQK-IDLSKDEVGKLTQIIQK 1064

RESULT 13
US-09-949-016-7404
; Sequence 7404, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7404
; LENGTH: 2047
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-7404

Query Match 2.0%; Score 184; DB 4; Length 2047;
Best Local Similarity 18.7%; Pred. No. 2.5e-06;
Matches 233; Conservative 195; Mismatches 453; Indels 366; Gaps 53;

Qy 209 BEKELPPELQIREKNERYCVLFDNDEHSHYDHYVLSQALDCELABAQLHTAIDKEGR 268
Db 147 BEKHEASEIQIKQOOSTSYRNQL--QKEVEISH-LKARQIALQDQLLKQ-----SAA 196
Qy 269 RAVKAGAYAAQCAEKEDIKSHSNVSHQ--LHVEVL-HSEIMAHQKFAELGSMWNKIM 325
```

```
Db 197 QSPSGAGVPATTAS---SSPAYGISHPSAFHDDMDPGDIISQOQENLSNEVSRL 253
Qy 326 SYSSDPFOIFQOACLRPEPDSNCLISRLMLDAKLYGARKILHELIPSPFFMEMEYK 385
Db 254 SEVGHWRHI-AQTSKAQGTDSQSEICKL-----282
Qy 386 KLFAMEFVKYKQLOKQEYISDHDORSISITALSVQMFTPTLARHLIEQONVISVITETL 445
Db 283 ----QNIIEKELKONRSQEI-DDHQHEMSVLQNAHQOKLTEISRRHRELSDY-----BER 332
Qy 446 LEVLPYLDNRNNKFNFGYSQDKGRVYAVICDLKYLILISKPTIWEHLRQWFLGFSRSP 505
Db 333 IEELLENLIQGG---SGVIETDLSKIYEMQKTIQVLQIEK--VESITKMEQLEDKIKDI 386
Qy 506 LKILTCMQMEIEIRQVQCHIEVDPDWEAAITAIQMLAKNILLMPQWACACDELLLVAYK 565
Db 387 NKCLSSAENDRDIIRREOEQLNVEK-----RQIM-----E 416
Qy 566 ECHKAVMRCSS--TSFISSTVWQSCGHSLETYSRVSEDLVSIH--LPLSRTLGLHV 620
Db 417 ECENLKSCSKLOPSAVKQSDTM-----TEKERILAQSASVEEVRLOQALSDAEN 467
Qy 621 RLSRLGAVSRHLEFVSPEDFQVEVLVEYPLRCLVLVAQVAVMWRNGLSLISQVFFYQ- 679
Db 468 EIMRLSSLNODNSLAE-DNLKLMRIEVLKEKESLLSQEKEEL---QMSLLKLNVEYEV 522
Qy 680 -----DVKCREBYDKDIIMLOIGASIMDPNKFEL-----LLVLQRYEL---AFAFNKT 724
Db 523 TKSTATRDISLDSLEHD---LRLNLEAKEBELNOSISEKETLIAIEBELDQNEQATKHM 579
Qy 725 ISTDQODLIKOVN---TLIEBMLQVLIYIGERYVPGVGNVTKEVTMREIHLICISPM 781
Db 580 ILIKQO--LSKQONEGDSIISKLKQDL-----NDEKKEV-----HQLEDDKM 619
Qy 782 PHSAIAKNLPENNETGLENVINKVATPKPGVSGHGYELKDBSLKDFNMVFFHYSKT 841
Db 620 D---ITKELDVQKEKLIQSEVALNDLHLTKQK-----LEDKVENLVD-----QLNKS 663
Qy 842 QHS-----KAHMOKKRRKQENKDEALPPPPPPFCFAPSKVINLNCIDMMVILTVER 897
Db 664 QESNVSIQENLELKEHIRONEE-----LSRIRNEL-----MQ 697
Qy 898 AIDTDSNLWTEGMLQWAFHILALGALLEBEKQOQKAPK--EVTDFYTHKASRLGSSAMNI 955
Db 698 SLNODSNSNFKDTLLKEREAEVRNLKQNLSELEQLNENLKVAFDVKMENKLVACEDV 757
Qy 956 QMLLEK-LKGIPOLEGQKDMITWILQMPD-----TVKRLREKSCLIIVATTSGSESI 1005
Db 758 RHQLEBCLAGNQLSLEKNTIVETLKEBGEIEABLCKWAKLLBEANKYKTEIELSNA 817
Qy 1006 KN-----DBITH-----DKEKAERKRKAE-----AARLHQKIMAQM 1037
Db 818 RNLTNSALQLEHEHLIKLQKKDMEIAELKKNIEQMDTDHKETKDVLSSELEQKLTOL 877
Qy 1038 SALQKNPIETHKLMYDNTSEMPGK-----EDSIMEEES 1070
Db 878 INKKEIFIEKLRKERSKQLELDKYSQLRKNNEILRQTIEBKDRSLGSMKEENHLOEEL 937
Qy 1071 TPAVSDYSRIA--LGPKRGPSTVE--KEVLTCILQOEBOEVKIENNAMVLSACVOKSTAL 1126
Db 938 ERLREEQSTAPVADPKTLDSTVELASEVSLNTIKHELEBEIKHKQKIIEDQNSKQML 997
Qy 1127 TQHRGKPIELSGEALDPLFMDPDLAYGTYTSGCGHMHAVCQKYPEAVQLSSQQRHVD 1186
Db 998 LQ-----SLQEQKEMDEF-----RY-----QHEQMATHQ 1024
Qy 1187 LPDLSGEVLCPLCKSLCNTVPI-----IPLQPKINSENADALAQLLTLARWITVLARI 1243
Db 1025 LP-LEKDEBI-----KSLQKTIIEQIKTQLHEERQDIQDTSNDFQEB-----TKV 1067
Qy 1244 SGYNIRHAKGENPIPIFFNQGMGDSITLPHSILSGVSESIKYNSIKEMWILFATTIYR 1303
```


Best Local Similarity 19.4%; Pred. No. 7.9e-06;			
Matches 200; Conservative 160; Mismatches 344; Indels 329; Gaps 45;			
QY	167	EPGRAGTIKENRCPINEEVIQVARKIPFSVIKYVEMTWEEKELPPELQ-----IREK 222	
Db	488	EDTRVATVSEKSR-----IMELEKDLALRVQVAEARRR 521	
QY	223	NERYCYLVFENDEHSHYDHYISLQALDC-----ELAAQIHTTADKEGRRAVKAGA 275	
Db	522	LESNKPAGDVMSLSLLOEISSLOKLEVTTRDHOEITSLKEHFGAREETHQKEIKA-L 580	
QY	276	YAACO---EAKEDIKS-----HSENVQSHPLEVEIHSMAHQFALRLG-SMNKIMS 326	
Db	581	YTATEKLSKENSLSKLEHANKENSVDIALWKSLETAIAASHQQAMEELKVSPSKGLGT 640	
QY	327	YSDPRQIFCQ-----ACLRBPDSNPCLISRLMLWDAKLYKGARKILHELI 374	
Db	641	ETAEPFAELKTQIEKMRLDYQHEIENLQOQDSER----- 674	
QY	375	FSSPFMEYKYLPAHEFYKYKQLOKEVISDDHRSISITALSVMQFTVPTLABHLIEE 434	
Db	675	-AAHAKEMALRAKLMKVIK-EKENSLEAIRSKLDAED-----OHLVEM 717	
QY	435	QNVISVITET---LLEVLPEYLDNRNKP--NFQGY---SODKLGVRVAVICDLKYILI 484	
Db	718	EDTLNKLQAEAIKVKELEVLQAKNEQTKVDINFTSOLKATEEKL-----LDL----- 765	
QY	485	SKPTIWTIRLMQFLGFRSPILKILTCMQMBEIRRVQGOHIEVDPDWEA--AIAIOMQL 542	
Db	766	-----DALRKASSEG-KSEMKCL--RQOLEAAEKQI-KHLEIEKNABESSKASSITREL 814	
QY	543	KNILLMFQWCACDEBLLIVAYKECHKAVMRCSTSPISSSKTVVQSCGSHSLETYSRVSE 602	
Db	815	QG-----RELKLTNLQE-----NLSEVSQVKETLEKELQILKEKFAEASE 854	
QY	603	DLVS IHLPLSRTIAGLHVRLSLGAVSLRHEFVSPEDFOVEVLVEYPLRCLVLVAQVVAE 662	
Db	855	EAVSVORSQOETVNLHQKEEQFNMLSS-----DLEKL----- 887	
QY	663	MWRNGLSLISQVFFYQD-----VKCREMYDKOI--IMLQIGASLMDPNKPELLLVLOR 714	
Db	888	---RENLADEAKFREKDEREQLIKAKAKEL-ENDIAEIMKMSGDNSSQLTK----- 935	
QY	715	YELAEAFNTI1STKQDOLIKQNTLIEEMLOVLIIYIGERYVPGVGNVTKEEVTMBREI1H 774	
Db	936	-----MNDLRLKRDV-----BELQLKLTKANEN--ASFLOKSIEDMTVK----- 974	
QY	775	LLCIEPMPHSAIAKNLPENENNETHLENVINKVATPKPGVGHGVYELKDESLSKDFNNY 834	
Db	975	---AEOSQOBAKGGHEEKKELEKRLKSLDEKKMET-----SHNQOELKAR----- 1017	
QY	835	FYHYSKTOHSAKHHMOKRRKQENKDEALPPPPPEFCPAPSKVINLNCDIMMYILRTV 894	
Db	1018	--YERATSETKTKEE-----ILQNL 1036	
QY	895	FERAITDNLNTEGMLQWAPHILALGILLEEKOLOKAPEEVTPFYHKASRLGSSANN 954	
Db	1037	QKTLTLDTEKLA--KGAREE-----NSGLLQLELEELAKQADKA-----KAAQTAEDAMO 1082	
QY	955	I--QMLEKLGIPQLEGOKDMITWILQMPDVTVRLREKSLIVATTSGSESIKNDEITH 1012	
Db	1083	IMEQMTKEKTETLASLEDTKQTNAKLQNELDTLKNNLNKNBELNKSKEILLTVENQKXEE 1142	
QY	1013	DKEKAERKGAEAARLHROKIMAQMSALQKNFIETHKLMYDNTSEMPGKEDSI-----ME 1067	
Db	1143	FRKEIETLKQAAQK-----SQLSALQE--ENVKL-----AEELGRSRDEVTSHQKLE 1189	
QY	1068	BESTPAVSDYR1ALGPKRGSPSTEKEVILTCILQEEQVKEIENNANVL1SACV-QKSTAL 1126	
Db	1190	EER-----SVLNNQLLEMKKRESKFIKA-----DEEKASLQKSI1TSAL1TEKDAEL 1238	
QY	1127	TOHRGKPIELSGE 1139	

Db 1239 EKLRNEVTVLGRG 1251

Search completed: July 9, 2005, 15:33:39
Job time : 58 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2005, 08:02:32 ; Search time 157 Seconds
(without alignments)
5704.627 Million cell updates/sec

Title: US-10-758-672A-2

Perfect score: 9224

Sequence: 1 MADEEAGGTERMEISAEPLQ.....EIASRQETNQLFGFNWQLL 1749

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 03:*

1: uniprot_sprot;*

2: uniprot_trembl;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9224	100.0	1749	2 Q8IWV7	Q8iw7 homo sapien
2	8928.5	96.8	1709	2 Q8IWY6	Q8iwy6 homo sapien
3	8617	93.4	1757	2 Q70481	Q70481 mus musculus
4	4327	46.9	1755	2 Q8IWV8	Q8iww8 homo sapien
5	4280	46.4	1755	2 Q6DIB9	Q6dib9 mus musculus
6	4278	46.4	1755	2 Q8CGW0	Q8cgw0 mus musculus
7	4262	46.2	849	2 Q8BN40	Q8bn40 mus musculus
8	4218.5	45.7	1756	2 Q6WKZ8	Q6wkz8 mus musculus
9	4209	45.6	811	2 Q96JY4	Q96jy4 homo sapien
10	3160	34.3	1275	2 Q15057	Q15057 homo sapien
11	3014.5	32.7	1246	2 Q8OU31	Q8ou31 mus musculus
12	2887.5	29.1	1109	2 Q8K216	Q8k216 mus musculus
13	2416.5	26.2	1679	2 Q7QSV6	Q7qev6 anopheles g
14	2280.5	24.7	1824	2 Q8SX71	Q8sx71 drosophila
15	2276.5	24.7	1824	2 Q9VX91	Q9vx91 drosophila
16	2248	24.4	446	2 Q792M3	Q792m3 mus musculus
17	2232.5	24.2	861	2 Q8BUL9	Q8bul9 m mus muscu
18	1945.5	21.1	1927	2 P91133	P91133 caenorhabdi
19	1782	19.3	777	2 Q8ROV7	Q8rov7 mus musculus
20	1716	18.6	333	2 Q6708	Q6708 homo sapien
21	1692	18.3	329	2 Q75492	Q75492 homo sapien
22	1301.5	14.1	575	2 Q6ZUD0	Q6zud0 homo sapien
23	1204	13.1	229	2 Q8CSK3	Q8csk3 mus musculus
24	1085	11.8	439	2 Q6P212	Q6p212 homo sapien
25	1038	11.3	1958	1 UBRI SCHPO	O60152 schizosacch
26	1020	11.1	2156	2 Q7SGT6	Q7sgt6 neurospora
27	1005.5	10.9	410	2 Q8R130	Q8r130 mus musculus
28	914	9.9	2052	1 UBRR SCHPO	O13731 schizosacch
29	898	9.7	1926	2 Q6BHH3	Q6bhh3 debaryomyce
30	848.5	9.2	1877	2 Q6CFU4	Q6cfu4 yarrowia li
31	709.5	7.7	1945	1 UBRI_KLULA	O60014 kluyveromyc

32 707 7.7 148 2 Q68DN9 Q68dn9 homo sapien
33 689.5 7.5 1950 1 UBRI_YEAST P19812 saccharomyc
34 686.5 7.4 1900 2 Q75F74 Q75f74 ashbya goess
35 679.5 7.4 1813 2 Q75F75 Q75f75 ashbya goess
36 661 7.2 1814 2 Q6BHH4 Q6bhh4 debaryomyce
37 639.5 6.9 1965 2 Q6PRD9 Q6prd9 candida gla
38 606.5 6.6 1225 2 Q9L295 Q9l295 arabidopsis
39 605.5 6.6 1787 2 Q6FW32 Q6fw32 candida gla
40 598 6.5 212 2 Q9H578 Q9h578 homo sapien
41 549.5 6.0 1872 2 Q07963 Q07963 saccharomyc
42 543.5 5.9 1709 2 Q6CJW9 Q6cjw9 kluyveromyc
43 541 5.9 2010 2 Q19330 Q19330 caenorhabdi
44 471.5 5.1 1869 2 Q86N11 Q86n11 caenorhabdi
45 470.5 5.1 1264 2 Q94JL2 Q94jl2 oryza sativ

ALIGNMENTS

RESULT 1

Q8IWV7 Q8IWV7 PRELIMINARY; PRT; 1749 AA.
AC Q8IWV7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Ubiquitin ligase E3 alpha-1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Han H.Q., Kwak K.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY061886; AAL32103.1; -
DR Genew; HGNC:16808; UBRI.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0006512; P:ubiquitin cycle; IEA.
DR InterPro; IPR000307; Ribosomal_S16.
DR InterPro; IPR003126; Znf_Nrscogmin.
DR Pfam; PF02207; zf-UBRI; 1.
DR SMART; SM00396; Znf UBRI; 1.
DR PROSITE; PS00732; RIBOSOMAL_S16; UNKNOWN_1.
DR KW Ligase.
SQ SEQUENCE 1749 AA; 200209 MW; 3AE0E1A749884971 CRC64;

Query Match 100.0%; Score 9224; DB 2; Length 1749;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1749; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MADEEAGGTERMEISAEPLQPTQRLASWDDQVDFYTAFLHLAQLVPEIYFAEMDPDL 60
DB 1 MADEEAGGTERMEISAEPLQPTQRLASWDDQVDFYTAFLHLAQLVPEIYFAEMDPDL 60
QY 61 KQESVQMSIFTPLEWYLFGEPPDICLEKXHSAGFQLCGRVFKSGETYSYSCDCAIDPT 120
DB 61 KQESVQMSIFTPLEWYLFGEPPDICLEKXHSAGFQLCGRVFKSGETYSYSCDCAIDPT 120
QY 121 CVLCMDCFQDSVHKHNRHYKMTSTGGGFCGDTAEWTKGPPFCVNHPEGRAGTIKNSRC 180
DB 121 CVLCMDCFQDSVHKHNRHYKMTSTGGGFCGDTAEWTKGPPFCVNHPEGRAGTIKNSRC 180
QY 181 PLNEEVIVQARKIPPSVIKYVEMTIWEKEKLPPELQIREKNERYCYLFNDEHSHYD 240
DB 181 PLNEEVIVQARKIPPSVIKYVEMTIWEKEKLPPELQIREKNERYCYLFNDEHSHYD 240
QY 241 VIYSQRLALDCELAQAHLHTTAIDKEGRAVRAGAYAAQCAEKEDIKSHSENVSQPLHV 300
DB 241 VIYSQRLALDCELAQAHLHTTAIDKEGRAVRAGAYAAQCAEKEDIKSHSENVSQPLHV 300
QY 301 EVLHSEIMAHQFALRLGLSGMNMKIMSYSDFRQIFQCAQLREPPDSENPCILSLMLWDA 360

Db 301 EVLHSEIMAHQFALRGLGSMWKNKIMSYSDFRQIFQACLRREPSNFCILSRMLMDA 360
QY 361 KLYGARKILHELIFSSPFMEYKYLKAFMEFVKYKQLOKEYISDHDHRSISITALSVO 420
Db 361 KLYGARKILHELIFSSPFMEYKYLKAFMEFVKYKQLOKEYISDHDHRSISITALSVO 420
QY 421 MTVTPTLARHLIEBQNVISVITETTLLEVLPYLDNRNKNFQYSQDKLGRVYAVICDLK 480
Db 421 MTVTPTLARHLIEBQNVISVITETTLLEVLPYLDNRNKNFQYSQDKLGRVYAVICDLK 480
QY 481 YILISKPTTWRLRMOPLRGRFSPLKILTCMQMBEIRROVGOHLEVPDMEAAIAIQM 540
Db 481 YILISKPTTWRLRMOPLRGRFSPLKILTCMQMBEIRROVGOHLEVPDMEAAIAIQM 540
QY 541 QLNKILLMFQEWACADELLLVAYKECHAVMRCSTSFISSTKTVVQSCGHSLETKSYRV 600
Db 541 QLNKILLMFQEWACADELLLVAYKECHAVMRCSTSFISSTKTVVQSCGHSLETKSYRV 600
QY 601 SEDLSVSIHLPLSRTLAGLHVRLSRGAVSRHLHFVSFEDPOVEVLVEYPLRCLVLVAQVY 660
Db 601 SEDLSVSIHLPLSRTLAGLHVRLSRGAVSRHLHFVSFEDPOVEVLVEYPLRCLVLVAQVY 660
QY 661 AEMWRNGLSLISQVYQDVKCRBMYDKIIMLQIGASLMDPNKFLLLVLQRYELAEA 720
Db 661 AEMWRNGLSLISQVYQDVKCRBMYDKIIMLQIGASLMDPNKFLLLVLQRYELAEA 720
QY 721 FNKTISTKQDQILKOYNTLIEEMQLVLIYVGERYVPGVGNVTKKEBVTWREIHLICIEP 780
Db 721 FNKTISTKQDQILKOYNTLIEEMQLVLIYVGERYVPGVGNVTKKEBVTWREIHLICIEP 780
QY 781 MPHSIAKVLNPNENNETHGLNVINKVATFKPGVSGHGVYELKDBSLKDFNMFYHYSK 840
Db 781 MPHSIAKVLNPNENNETHGLNVINKVATFKPGVSGHGVYELKDBSLKDFNMFYHYSK 840
QY 841 TQSHKAEHMKRKRKQENKDEALPPPPPEFCFAPSKVINLNCNDIMWILRTVFPRAID 900
Db 841 TQSHKAEHMKRKRKQENKDEALPPPPPEFCFAPSKVINLNCNDIMWILRTVFPRAID 900
QY 901 TDSNLWTEGMLQNAFHILALGLLEERQKQKAPPEEVEVTFDYHKASRLGSSAMNIQMLLE 960
Db 901 TDSNLWTEGMLQNAFHILALGLLEERQKQKAPPEEVEVTFDYHKASRLGSSAMNIQMLLE 960
QY 961 KLKGIPOLEGQKDMITWILQMPFTVRLREKSKLIVATTSGBSINKDEITHDKAEARK 1020
Db 961 KLKGIPOLEGQKDMITWILQMPFTVRLREKSKLIVATTSGBSINKDEITHDKAEARK 1020
QY 1021 RKAERARLHROKIMQMSALQKNPIETHKIWDNTSEMPGKBDIMEEESTPAVDSYRI 1080
Db 1021 RKAERARLHROKIMQMSALQKNPIETHKIWDNTSEMPGKBDIMEEESTPAVDSYRI 1080
QY 1081 ALGPKRGPSVTEKEVLTICILCOBQEVKINENAMVLSACVQKSTALTQHRGKPIELSGEA 1140
Db 1081 ALGPKRGPSVTEKEVLTICILCOBQEVKINENAMVLSACVQKSTALTQHRGKPIELSGEA 1140
QY 1141 LDPLFMDPDLAYTGTTCGSHVMHVCWKYFRAVOLSSQORIHVDLFDLESGEYLCPLC 1200
Db 1141 LDPLFMDPDLAYTGTTCGSHVMHVCWKYFRAVOLSSQORIHVDLFDLESGEYLCPLC 1200
QY 1201 KSLCNTVPIPIIPIQOKINSENADALAQLLTLARWITQTLARISGYNIRHAKENPIPIF 1260
Db 1201 KSLCNTVPIPIIPIQOKINSENADALAQLLTLARWITQTLARISGYNIRHAKENPIPIF 1260
QY 1261 FNQGMGDSITLFFHSILISFGVSESIKYSNSIKENWILFATTIYRIGLKVPPDERDPRVPM 1320
Db 1261 FNQGMGDSITLFFHSILISFGVSESIKYSNSIKENWILFATTIYRIGLKVPPDERDPRVPM 1320
QY 1321 TWSTCAFTIQAIEINLGDGKPLFGALONRQHNGKALMQFAVQAQITCQVLIQKHLVR 1380
Db 1321 TWSTCAFTIQAIEINLGDGKPLFGALONRQHNGKALMQFAVQAQITCQVLIQKHLVR 1380
QY 1381 LLSVLPNPKSDETPCLLSIDLPHVLGAVLAPPSLYWDDPVDLPQSSVSSSYNHLYLPH 1440

Db 1381 LLSVLPNPKSDETPCLLSIDLPHVLGAVLAPPSLYWDDPVDLPQSSVSSSYNHLYLPH 1440
QY 1441 LITMAHMLQILLTVDVTGLPLAQVQSDSEAHASASSFFAISIQYTSSTGSGCDIPGHWLWS 1500
Db 1441 LITMAHMLQILLTVDVTGLPLAQVQSDSEAHASASSFFAISIQYTSSTGSGCDIPGHWLWS 1500
QY 1501 LKNGITTPYLRCALPFFHYLLGVTTPPEELHTNSABGEYSALCSYLSLPTNLFLLFQBYWDT 1560
Db 1501 LKNGITTPYLRCALPFFHYLLGVTTPPEELHTNSABGEYSALCSYLSLPTNLFLLFQBYWDT 1560
QY 1561 VRPLLRQWCADPALLNCLKQKNTVVYPRKRNLSLIELPDDYSCLLNQASHFRCPSADDE 1620
Db 1561 VRPLLRQWCADPALLNCLKQKNTVVYPRKRNLSLIELPDDYSCLLNQASHFRCPSADDE 1620
QY 1621 RKHPVLCLFCGAILCSQNICCOBIVNGEVEGACIFALHCGAGVCIFLKIRCRVVLVEG 1680
Db 1621 RKHPVLCLFCGAILCSQNICCOBIVNGEVEGACIFALHCGAGVCIFLKIRCRVVLVEG 1680
QY 1681 KARGCAYPAPYLDYGETDPGLKRGKGNPLHLRERKRLHLVWQOHCIIIEIARSQETNOM 1740
Db 1681 KARGCAYPAPYLDYGETDPGLKRGKGNPLHLRERKRLHLVWQOHCIIIEIARSQETNOM 1740
QY 1741 LFGFNWQLL 1749
Db 1741 LFGFNWQLL 1749

RESULT 2
QBIWY6 PRELIMINARY; PRT; 1709 AA.
AC Q8IYW6; PRELIMINARY; PRT; 1709 AA.
DT 01-MAR-2003 (TREMELrel. 23, Created)
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
DE 01-MAR-2004 (TREMELrel. 26, Last annotation update)
DE UBRI E3a ligase (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22340442; PubMed=12434312;
RA Dgany O., Avidan N., Delaunay J., Kraenow T., Shalmon L., Shalev H.,
RA Eidelitz-Markus T., Kapelushnik J., Cattani D., Pariente A.,
RA Fullez M., Cretien A., Schischmanoff P.O., Iolascon A., Fibach E.,
RA Koren A., Rossier J., Le Merrier M., Yaniv I., Zaizov R., Ben-Asher E.,
RA Olender T., Lancet D., Beckmann J.S., Tamary H.;
RT "Congenital dyserythropoietic anemia type I is caused by mutations in
RT codanin-1";
RL Am. J. Hum. Genet. 71:1467-1474(2002).
DR EMBL; AF525401; AAO14997.1; --.
DR GO; GO:0016874; F:ligase activity; IEA.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
DR GO; GO:0006512; P:ubiquitin cycle; IEA.
DR InterPro; IPR000307; Ribosomal_S16.
DR InterPro; IPR003126; Znf_Nrecogin.
DR Pfam; PF02207; zf-UBR1; 1.
DR SMART; SM00396; ZNF_UBR1; 1.
DR PROSITE; PS00732; RIBOSOMAL_S16; UNKNOWN_1.
KW Ligase.
FT NON_TER 1 1709 1709
FT SEQUENCE 1709 AA; 195278 MW; B2B17231A1020FID CRC64;
Query Match 96.8%; Score 8928.5; DB 2; Length 1709;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1697; Conservative 1; Mismatches 11; Indels 1; Gaps 1;
QY 18 LPQTQORLASWDDQVDFYTAFLHLAQLVPELYFAEMDPDLKQESVQMSIFTPLEWY 77
Db 1 LPQTQORLASWDDQVDFYTAFLHLAQLVPELYFAEMDPDLKQESVQMSIFTPLEWY 60
QY 78 LFGEDPDICLEKLKHSQAFQCGRVFKSGRTTYSRCDCAIDPTCVLCMDCFQDSVHKHNR 137

Db 61 LFGSDPDI CLEKUKHSGAFQICGRVFKSGETTYSRCDAIDPTCVLCMDCFQSVHKNR 120
QY 138 YKHHTSTGGGFCDCGDTAEAKTGPFCVNHPEPGRAGTIKENSRCPLNEEVIVQARKIPPSV 197
Db 121 YKHHTSTGGGFCDCGDTAEAKTGPFCVNHPEPGRAGTIKENSRCPLNEEVIVQARKIPPSV 180
QY 198 IKVVEMTIWEEBKELPPELOIREKNERYCYVLFNDEHSHYDVIYLSQALDCELAQAQ 257
Db 181 IKVVEMTIWEEBKELPPELOIREKNERYCYVLFNDEHSHYDVIYLSQALDCELAQAQ 240
QY 258 LHATAIDKGERRAVAKAGAYAAQCAEKEDIKSHSENSVSHPLHVEVLHSEIMAKQKFAIRL 317
Db 241 LHATAIDKGERRAVAKAGAYAAQCAEKEDIKSHSENSVSHPLHVEVLHSEIMAKQKFAIRL 300
QY 318 GSWNKMYSYSSDFRQIFQACLRBEPDSENCLISRLMLWDAKLYKGARKILHELIFSS 377
Db 301 GSWNKMYSYSSDFRQIFQACLRBEPDSENCLISRLMLWDAKLYKGARKILHELIFSS 360
QY 378 FFMEMEYKCLPAMEFVKYKQLOKEYSIDDDHDSISITALSVMQFTVPTLARHLIBEONV 437
Db 361 FFMEMEYKCLPAMEFVKYKQLOKEYSIDDDHDSISITALSVMQFTVPTLARHLIBEONV 420
QY 438 ISVITETLLEVLPEYLDNRNKNFQGYSDKLGVRVAVICDLKYLILISKPTIINTERLRMQ 497
Db 421 ISVITETLLEVLPEYLDNRNKNFQGYSDKLGVRVAVICDLKYLILISKPTIINTERLRMQ 480
QY 498 FLGFRSFLKILTCMQGMEIRQVGHIEVDPDWEAAIAIQMLKNILMLFQEWCAQDE 557
Db 481 FLGFRSFLKILTCMQGMEIRQVGHIEVDPDWEAAIAIQMLKNILMLFQEWCAQDE 540
QY 558 ELLLVAYKECHKAVMRCSTSFISSTKTVQSGHSLSTKSYRVSDELVSJHLPLSRTLAG 617
Db 541 ELLLVAYKECHKAVMRCSTSFISSTKTVQSGHSLSTKSYRVSDELVSJHLPLSRTLAG 600
QY 618 LHVRLRLGAVSRHLHREFVSPEDQVVLVEYPLRCLVLAQVVAEWMRRNGLSLISQVY 677
Db 601 LHVRLRLGAVSRHLHREFVSPEDQVVLVEYPLRCLVLAQVVAEWMRRNGLSLISQVY 660
QY 678 YQDVKREEMMYDXDIIMLQIGASLMDPNKFLLVQRYELAEAFNKTISTKQODLQKYN 737
Db 661 YQDVKREEMMYDXDIIMLQIGASLMDPNKFLLVQRYELAEAFNKTISTKQODLQKYN 720
QY 738 TLIEMLQVLIYVGRYVPGVGNVTKGVTMBEIIHLLCIRPMPSHSAIAKNIPENENNE 797
Db 721 TLIEMLQVLIYVGRYVPGVGNVTKGVTMBEIIHLLCIRPMPSHSAIAKNIPENENNE 780
QY 798 TGLENVINKVATPKKQVSGHGVYELKDBSLKDPNMFYHYSTQHSKAEHMOKKRRQOE 857
Db 781 TGLENVINKVATPKKQVSGHGVYELKDBSLKDPNMFYHYSTQHSKAEHMOKKRRQOE 840
QY 858 NKDEALPPPPPPFCFAPSKVINLLNCDIMMYILRTVFERAIDTDSNLMWTEGMLQWAFHI 917
Db 841 NKDEALPPPPPPFCFAPSKVINLLNCDIMMYILRTVFERAIDTDSNLMWTEGMLQWAFHI 900
QY 918 LALGLLEEKQOLQKAPPEEVTPDYFKASRLGSSANNIOMLKLKXGIPQLEQKDMITW 977
Db 901 LALGLLEEKQOLQKAPPEEVTPDYFKASRLGSSANNIOMLKLKXGIPQLEQKDMITW 960
QY 978 ILQMPFTVKRLREKSLIVATTSGSESINKDEITHDKEAERKKAERAAALHROKTMQOM 1037
Db 961 ILQMPFTVKRLREKSLIVATTSGSESINKDEITHDKEAERKKAERAAALHROKTMQOM 1020
QY 1038 SALQKQFIETHKLWYDNTSMPCKEDSIMEESTPAVSDYSRIALGPKGSPVTEKEVL 1097
Db 1021 SALQKQFIETHKLWYDNTSMPCKEDSIMEESTPAVSDYSRIALGPKGSPVTEKEVL 1080
QY 1098 CILCOBEQEVKIENNAWLSACVQKSTALTQHRGKPIELSGEALDPLFMDPDLAYGTYTG 1157
Db 1081 CILCOBEQEVKIENNAWLSACVQKSTALTQHRGKPIELSGEALDPLFMDPDLAYGTYTG 1140
QY 1158 SCGHVMAHVCWKYFEAVQLSSOORIHVDLFDLESGEYLCPLCKSLCNTVPIPILOPOK 1217

Db 1141 SCGHVMAHVCWKYFEAVQLSSOORIHVDLFDLESGEYLCPLCKSLCNTVPIPILOPOK 1200
QY 1218 INSENADALAQLLTLARWIQTVLARISGVNIRHAKGENPIPIFFNQGMDSTLEPHSILS 1277
Db 1201 INSENADALAQLLTLARWIQTVLARISGVNIRHAKGENPIPIFFNQGMDSTLEPHSILS 1260
QY 1278 FGVESSEIKYNSIKEMVILLPATYIYRIGLVKVPDERDPRVPMLTWSTCAFTIOAIENLLG 1337
Db 1261 FGVESSEIKYNSIKEMVILLPATYIYRIGLVKVPDERDPRVPMLTWSTCAFTIOAIENLLG 1320
QY 1338 DECKPLFGALQNRQHNGLKALMQFAVAQRITCQVLIQKHLVRLLSVVLNPKISEDTPL 1397
Db 1321 DECKPLFGALQNRQHNGLKALMQFAVAQRITCQVLIQKHLVRLLSVVLNPKISEDTPL 1380
QY 1398 LSIDLPHVLVGAVALPSPSYWDDPVDLPQSSVSSSSNNHLYLPHLITMAHMLQILLTVDTG 1457
Db 1381 LSIDLPHVLVGAVALPSPSYWDDPVDLPQSSVSSSSNNHLYLPHLITMAHMLQILLTVDTG 1440
QY 1458 LPLAQVQDESEHAHSSAFBAISQYTSIGSDIPGWYLVWSLXNGITPYLRCAALFFH 1517
Db 1441 LPLAQVQDESEHAHSSAFBAISQYTSIGSDIPGWYLVWSLXNGITPYLRCAALFFH 1500
QY 1518 YLLGVTPPELHTNSAEGYSALCSYLSLPTNLFLFQEWYDTPRPLQRCWADPALLNC 1577
Db 1501 YLLGVTPPELHTNSAEGYSALCSYLSLPTNLFLFQEWYDTPRPLQRCWADPALLNC 1560
QY 1578 LKQKNTVVRYPRKRNLSIELPDDYSCLLQASHFRCPRSADDERKHPVLCFCGAILCSQ 1637
Db 1561 LKQKNTVVRYPRKRNLSIELPDDYSCLLQASHFRCPRSADDERKHPVLCFCGAILCSQ 1620
QY 1638 NTCQBIQVNGEEVGACIFHALHCGAGVCIFLKIRECRVVLVEGKARGCAYPAPYLDYGE 1697
Db 1621 NTCQBIQVNGEEVGACIFHALHCGAGVCIFLKIRECRVVLVEGKARGCAYPAPYLDYGE 1680
QY 1698 TDPGLKRGKGNPLHLRERYRKLHLVWQOHC 1727
Db 1681 TDPGLKRGKGNPLHLRERYRKLHLVWQOHC 1709
RESULT 3
070481
ID 070481 PRELIMINARY; PRT; 1757 AA.
AC 070481;
DT 01-AUG-1998 (TREMELrel. 07, Created)
DT 01-AUG-1998 (TREMELrel. 07, Last sequence update)
DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)
DE Ubiquitin-protein ligase E3-alpha.
GN Names=Ubr1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98318563; PubMed=9653112; DOI=10.1073/pnas.95.14.7898;
RA Kwon Y.T., Reiss Y., Fried V.A., Hershko A., Yoon J.K., Gonda D.K.,
RA Sangan P., Copeland N.G., Jenkins N.A., Varshavsky A.;
RT "The mouse and human genes encoding the recognition component of the
RT N-end rule pathway."
RL Proc. Natl. Acad. Sci. U.S.A. 95:7898-7903(1998).
DR EMBL; AF061555; AAC40165.1; -.
DR FIR; T14318; T14318.
DR MGD; MGI:1277977; Ubr1.
DR GO; GO:000151; C:ubiquitin ligase complex; IGI.
DR GO; GO:0005515; F:protein binding; IGI.
DR GO; GO:0004840; F:ubiquitin conjugating enzyme activity; IGI.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IGI.
DR GO; GO:0006512; P:ubiquitin cycle; IGI.
DR GO; GO:0006511; P:ubiquitin-dependent protein catabolism; IGI.
DR InterPro; IPR000307; Ribosomal_S16.
DR InterPro; IPR0003126; Znf_Nrecognin.
DR Pfam; PF02207; zf-Ubr1; 1.
DR SMART; SM00396; Znf_Ubr1; 1.

DR Genew; HGNC:21289; C:ubiquitin ligase complex; IEA.
 DR GO:000151; C:ubiquitin ligase complex; IEA.
 DR GO:0016874; F:ligase activity; IEA.
 DR GO:0004842; F:ubiquitin-protein ligase activity; IEA.
 DR GO:0008270; F:zinc ion binding; IEA.
 DR GO:0016567; P:protein ubiquitination; IEA.
 DR InterPro; IPR002345; Lipocalin.
 DR InterPro; IPR00408; Reg chr condens.
 DR InterPro; IPR003126; Znf Nrecognin.
 DR InterPro; IPR001841; Znf ring.
 DR Pfam; PFO2207; zf-UBR1; 1.
 DR SMART; SM00184; RING; 1.
 DR SMART; SM00396; Znf UBR1; 1.
 DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_1.
 DR PROSITE; PS00626; KCC1_2; UNKNOWN_1.
 DR Ligase.
 SQ SEQUENCE 1755 AA; 200536 MW; 04B14FCB13E21808 CRC64;

Query Match 46.98; Score 4327; DB 2; Length 1755;
 Best Local Similarity 47.28; Pred. No. 4.9e-245;
 Matches 841; Conservative 339; Mismatches 543; Indels 58; Gaps 27;

QY 1 MADE---EAGGTER--MEISAEPLQTPQRLASWDDQVDFTYAPLHLAQLVPEIYFAEM 55
 DB 1 MASELEPEVQADRSLECSAE-----EIAQWLQATDLTREYQHLAHYVFKIYCRGP 54

QY 56 DPDLKQESVQMSIFPTLEWYLFGEEDPDICLEKLKHSQ-AFOLCGRVKSGTETYSRCD 114
 DB 55 NPPQKEDMLAQHVLLGPMWYLCGEDPAPFGPKLEQANKPSHLCKRVPKVGEPYTSRCD 114

QY 115 CATDPTCVLCMDCFQDSVHNKHYKMTSFGGFCDCGDTGTEAKWTGPPCVNHFPRAG-T 173
 DB 115 CAVDPTCVLCMECFGLSIHRDHYRMTSGGFCDCGDTGTEAKWKEGPPYQKHEINTSEIE 174

QY 174 IKNSRCPLNEEVIQARKLFPSPVIVVEMTWEEKEKLPPLELOIREKNERYCYVLEND 233
 DB 175 EEDPLVHLSUEDVIARTYNIPTATRYAVAILTWEKESELPADELVKESDITYCMLEND 234

QY 234 EHSYDRVYLSQALDCELAELHTTAIDKEGRBAVKAGAYAAQCEAKEDIKSHSENV 293
 DB 235 EVHTYEQVYTLQKAVNCTQKEAIGPATTVDRDGRSVRYGDFQYCEQAKSVIVRNTSRQ 294

QY 294 SQHPLHVEVLHSHMAHQFALRLGSMNKIMSYSDFRQIFQACLRPEEDSENCLIS 353
 DB 295 TK-PLKVQVWHSHIVAHQFGLKLLSLGSIIGYSDGLRRILCQVGLQEGPDGENSESLVD 353

QY 354 RLMLDAKLYGARKILHELIFSSPFMEWEYKGLFAMEFVKYKOLQKEYISDDHDSIS 413
 DB 354 RLMSDKLWKGARSVTHQLFMSLLMDLKYKLLFAVRFAKNTQQIQDFMEDDHERAVS 413

QY 414 ITALSVMQFTVPTLARLHIEEQNVISVITETLELPEYLDNRNKNFQCYG---QDKLG 470
 DB 414 VTALSQVFPPTAPTLARMLITEENLMSIINKTFMDHL-RHRDAQRFQERYTALQAFKR 472

QY 471 RVAVATCDKYILISKPTIWTRELKMOLEGRFSFLKILTCMOGMEIERRQVQCHIEVDP 530
 DB 473 RVQSLILDYVLISKPTESDRLQKFLQEGDARLELLKCMQMDPITRQVQCHIEVDP 532

QY 531 DWEAATAIQMLKNILLMEQWACDEBELLVAYKECHKAVMECSFISSTKTVQVS-C 589
 DB 533 EWEAATLQMLKTHVISMMDWCASDEKLEIAYKKCLAVLMQCHGYTDGEOPIITLSIC 592

QY 590 GHSLETKSVYRSDLSIHLPLSRLTAGLHVRLSLGAVSRHLHFVSFEDFQVEVLVRYVP 649
 DB 593 GHSVETIRCVSQEKVSIHLPVSRLLAGLHVLSKSEYAKFPPELLPLSELSPMLIEHP 652

QY 650 LRCLVLVAQVVAEMWRNGLSLISQVFFYODVKCREEMVDKDIIMLQIGASLMDPNKFL 709
 DB 653 LRCLVLCAQVHAGWRNGLSLNQIYVYVHNKCREEMFDKDVVMLQTVGSMMDPNHFLM 712

QY 710 LVLYORVELAEAFN-----KTIISTK--DQDLIKQYNTLIEMLQVLYIYGERVYVGVNV 762
 DB 713 IMLSRPELQIFSTPDYGRFSSSEITHKDVVQNNLIEMLYLIIMLVYGFSPGVGV 772

RESULT 5

Q6DIB9

PRELIMINARY;

PRT; 1755 AA.

QY 763 TKEVTVMRBIHLLCTEPMHPSAIAKNLPENNENNETGLENNVINKVATFKKPGVSGHGYE 822
 DB 773 NATDEIKREBIHQLSIKPMHSELVSLPEDENKETCMESVIEAVAFKKPKGLTGRGMYE 832

QY 823 LKDESLSKDFNMVYHYKTOHSAEKHOKKRRKQENKDEALPPPPPEFCFAPSKVNL 882
 DB 833 LKPECAKBNLYFYHFRABEQSKAEBAQRKLKQNRREDTALPPVLPFPFCLFASLVN 892

QY 883 NCDIMMYILRTVFERAIDTDSNLWTEGLMQAPHILALGLLEEKQOLQKAPBEE-VTFDF 941
 DB 893 QSDVMLCINGTILQWAVEHNGYAVSSEMLQRLVHLIGMALQEBQKHLNTEHVVTFTF 952

QY 942 YHKASRLGSSAMN---IQMLLEKLGIPQLEGQDMITWILQWDTVTKRLREKS-CLIVA 997
 DB 953 TQKISKPGAPKNPSILAMLETQNPYLEVHKDMIRWLKTFNVAVKQMBRESSPTSPA 1012

QY 998 TTSGESIKNDREITHDKBAERKKAARLHRKQKIMQASALQKQKFIETHKLMYNTSE 1057
 DB 1013 ETEGTM---ESSRDKDAERKKAARLHRKQKIMQASALQKQKFIETHKLMYNTSE 1069

QY 1058 MPFKEDSIMBEEESTPAVSYSRIALGPKRGPSVTEKEVLTLCILQEEQEVKIENNAVL 1117
 DB 1070 LDASTSAVLDH--SPVASDWTLTALGPAQTQVPEQRQFTVCILQEEQEVKVESRAVLA 1127

QY 1118 ACVQKSTALTQHRGKPIELSGEALDPLFMDPLAYOTYTTGSCGHVMAVCWKYFRAVL 1177
 DB 1128 AFVQRSTVLSKNRSKFIQ-DPEKYDPLFMDPLSCGTHSSCGHIMHAHCWQRYFDSVQA 1186

QY 1178 SSQQ-----RIHVDLPDESSEVCLPCKSLGNTVPIPILOPKINSENADALAQLLT 1232
 DB 1187 KQRRQORLRLHTS-YDVENGEPFLCPLCBLCLNTVPLIL-LPPNPFNNRLN-FSQPNL 1243

QY 1233 ARWQITVLARIISGYNIRHAKGNPIPIPFNQMGDSTLBFHSTLSFGVSESSIKYSIKE 1292
 DB 1244 TQWIRTSIQIKALQPLRKEESTPNNASTKNSENVDELQLPQEGFRDPKPIYSSIKE 1303

QY 1293 MVLPATTIYRIGLVKPPDERPVPMLTWSTCAFTIOIENLLGDEGKPLFGALQNRQH 1352
 DB 1304 MLTFTGATYKVLKHPNEEDPRVIMCWGSCAYTIQSIERTLSDEKPLFGPLPCRLD 1363

QY 1353 NGKLALMQFAVQRTCPQVLTQKHLVRLSVLVNLIKSEDTPECLISIDLPHLVGNVLA 1412
 DB 1364 DCLRSITRFAAAHWTVASVSVVQGHFCKLPASLVPNDSHEELFCILDDIMFHLLVGLVA 1423

QY 1413 FPSLYWDDPVDLPQSSVSSYNHLYFLHITWAHMLQILL---TVDGLPLAQVQDSEE 1469
 DB 1424 FPALQCD-----FSGISLGTGDLHIFHLVTAHIIQILLTSTCEENGMD--QENPCEE 1476

QY 1470 AHSASSFFABISQYTSGSTGCDIP-GWYLVSLKNGITPYLRCAALFFHYLLGVTPPEEL 1528
 DB 1477 ESAVLALYKTLHQYT-GSALKBIPSGWLWRSVRAGIMPFLKCSALFFHYLVGVPSPDI 1535

QY 1529 HTNSAGEVSALCSYLSLPTNLFLPQBYWDTVRLPQWCADPALLNCLKQNTVVRYP 1588
 DB 1536 QV-PGTSHPHELCYLSLNNLCLFQENSEIMNSLIESWCNRSSEVKRYLEGERDAIRY 1594

QY 1589 RKNSLIELDDYSCLLNOASHFRCPRSADDERKHPVLCFCGAILLCSONICQOETVNGR 1648
 DB 1595 RESNKLINLPEDYSSLINQASNFCPKSGGDKSRAPTLCLVCSGLCSYCCQTELEGE 1654

QY 1649 EYVACIFHALHCGAGVCIPLKIRECRVLVVEGKARGCAYPAPYLDYEGYTDGPKRGK 1708
 DB 1655 DVGACTAHTYSCSGVGIFLVRVCEQVLQAGTKGCFYSPVLDYDGETDQGLRGK 1714

QY 1709 HLSRERYKHLVWQOHCIIIEIARSQETNQMFGFNWOLL 1749
 DB 1715 HLCKERFKIKQKLMHQHSVTEBIGHAQEANTQVLVGDIMQHL 1755

QY 1339 EGKELGALQNRHNGIKALMQFAVAQRITCPQVLTQKHLVRLSVVLPHNIKSEDTFCLL 1398
 DB 1350 EKEPVFGPLCRLLDDCLRLTRFAAAHRTVALLPVVQGHFKCLFASLVPSDSYEDLFCIL 1409
 QY 1399 SIDLFVLGVALFSLYWDPDVLPQSSVSSYNHLYLFLHITWAHMLQILL---TVD 1455
 DB 1410 DIDMFLLVGLVAFALQCD-----FSGSSLATGDLHIFHLVTHAHIVQIILLTCTEE 1464
 QY 1456 TGLPLAQVQBSDEAHASFPFABIQYTSGSGCIDIPGWYLVSLKNGITPYLRCAALF 1515
 DB 1465 NGMD--QENPTGEELAILSLHKTQYTGSALEKAPSGWHLWRSVRAAIMPLFKSALF 1522
 QY 1516 FHYLLGVTPEELHTNSAEYCALCSYLSLPTNLPLLFQBYWDTVRPLLRWCADPALL 1575
 DB 1523 FHYLNGVAPDPLQV-SGTSHFHLNLYSLPTNLHLFOENSIMNLSIESCQSEVK 1581
 QY 1576 NCLQKQNTVRYPRKNSLITELPDDYSCILNQASHFCRPSADDERKHVPVLCFCGAILC 1635
 DB 1582 RYLANGGALSYPGANKLIDLPEDYSSLIQNASNFCPSGSGKSRAPTLCLVCSLLC 1641
 QY 1636 SONICCOEIVNGBEVGACIFALHCGAGVCIFLKIRECRVVLVEGKARGCAYPAPYLDEY 1695
 DB 1642 SQSYCCQAELEGEDVGACTAHTYSCSGAGIFLRVRECQVFLAGTKGCFYSPPLDDY 1701
 QY 1696 GETDPGLKGNPLHLRERYRKLHLVQOHCITIERARSQETQMFLGFNWOLL 1749
 DB 1702 GETDQGLRRGNPLHLCOERFRKIQKLWQOHSITEBIGHAQEANTVLGIDWOHL 1755

RESULT 6

Q8CGW0 PRELIMINARY; PRT; 1755 AA.
 AC Q8CGW0;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Ubiquitin ligase E3 alpha-II (Fragment).
 GN Name=E130209G04Rik;
 OS Mus musculus (Mouse);
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C3H/HeN;
 RA Han H.Q., Kwak K.;
 RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY061885; AL32102.1; -;
 DR MGD; MGI:1925978; E130209G04Rik.
 DR GO; GO:0005515; P:protein binding; IPI.
 DR InterPro; IPR002345; Lipocalin.
 DR InterPro; IPR000408; Reg chr condens.
 DR InterPro; IPR003126; ZnF_Nrcognin.
 DR InterPro; IPR001841; ZnF_ring.
 DR Pfam; PF02207; zf-UBR1; 1.
 DR SMART; SM00184; RING; 1.
 DR SMART; SM00396; ZnF_UBR1; 1.
 DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_1.
 DR PROSITE; PS00626; RCC1_2; UNKNOWN_1.
 KW Ligase.
 FT NON_TER 1755 1755
 SQ SEQUENCE 1755 AA; 199164 MW; 21DD9DEB8281B39C CRC64;

Query Match 46.4%; Score 4278; DB 2; Length 1755;
 Best Local Similarity 46.5%; Pred. No. 3.7e-242; Indels 84; Gaps 28;
 Matches 835; Conservative 341; Mismatches 534;
 QY 1 MADE---EAGGTER--MEISAEPLQTPQRLASWDDQVDYFATFLHLAQLVPEIYFAEM 55
 DB 1 MASEMEPEVQADRSLECSAB-----EIAGRWLQATDLNREVYQHLAHCVPKIYCRGP 54
 QY 56 DPDLKQBSVQMSITPFLWYFGEDPDICLEKLKHSQ-AFOLGCRVPKSGETTVSCRD 114
 DB 1129 FVQRSTVLSKORTKI-ADPEKYDPLFMHPDLSGTHTSCGHHVMAHCWQRYFDSVQAK 1187

DB 55 NPPFPQKEDTLAQHILGPMWETICADPALGPPKLEQANKPSHLGCRVPKVGEPPTVSCRD 114
 QY 115 CAIDPTCVLCMDCFQDSVHKNHRYKHTSTGGFCDCGDTAEWKTGPFVFNHPPGAGTI 174
 DB 115 CAVDPTCVLCMEFLGSIHRDHRYRMTTSGGFCDCGDTAEWKEGYPQCKHKLSSSEVV 174
 QY 175 -KENSRCPLNEBIVIOARKIPSVIKYVYVEMTIWEEKELPPELQIUREKNERYCVLFPND 233
 DB 175 EEBEDPLVHLSVDVIARTYNIPAFMFRYAYDILTWKESELPEDLREVAEKSDTYTCHLFPND 234
 QY 234 EHSYDHYVLSQRLADCELAQLHTTAIDKEGRAVKAGAYAAOCEAKEDIKSHSENV 293
 DB 235 EHVITEQVYITLOKAVNCTQKEAIGPATVDRGRPVRYGDFQYCOQAKTVJVRMSRQ 294
 QY 294 SOHPLHVEVLSHSEIMAHQKFPALRLGSMNKNISYSSDFRQIFQACLRBEPDSENCLIS 353
 DB 295 TK-PLKQVQVHSSVAAHQNFGLKALSWLGSVIGSYDGLRILCOVLQSGPDGENSESLVD 353
 QY 354 RLMLWDKLYKGARKILHELIFFSSPFMEYKYLPAWPFVYKYQLOKEYISDDHDSIS 413
 DB 354 RLMLNDSKLMKARSVYHQLFMSSLLMDLYKCLFALRFAKNYRQLQDFMEDDHERAVS 413
 QY 414 ITALSVQMTPTVPTLARHLIEQNVISVITETILLEVLPEYLDNRNKNFQGYG--QDKLG 470
 DB 414 VTALSVQFTTATILARMLTEENLMTVIIKAFMDHL-KHRDAQRFQFERYTALQAFKR 472
 QY 471 RYVAVICDLKYILISKPTTWTBLRMQFLEGFRSFLKILTCMQGMBEIRRQVGHIEVDP 530
 DB 473 RVQSLILDLYVLSKPTWSDLRQKFLQGFDAFLLELLKCMQGMDDPITRQVQGHIEHP 532
 QY 531 DWEEAIAIQMLKNILLMOEWCADEBELLVAYKECHKAVMRCSTSPISSTKTVVQS-C 589
 DB 533 EWEAAFTLQMKLTHVISWQVDCALDEKVLIEAYKCLAVLTQCHGGFTDGEOPITLSIC 592
 QY 590 GHSLETKSVRSSEDLVS IHLPLSRTLAGHLVRLSLGAVSRLSEHFSPDPQVEVLVEYP 649
 DB 593 GHSVETIRYCVSQEKVS IHLPLSRLLAGHLVRLSEVAYKPELLPLSELSPMPLIEHP 652
 QY 650 LRLCLVLAQVAEMRRNGLSLISQVYFYQVCKREMYDKDIIMLQIGASLMDPNKFL 709
 DB 653 LRLCLVLAQVHAGWRNGFSLVNIQIYHYHVKCRREMPDKDIIMLQVGSWMDPNHFLM 712
 QY 710 LVLYRYELAEAPN-----KTIISTK--DQDLIKQYNTLIBEMLOVLIYIYGERVPGVNV 762
 DB 713 IMLSRFELYQLFSTPDYGRFSSEVTHKDVQVQNNLIBEMLYLIIMLVGERFNPVGQV 772
 QY 763 TXBEVYTMREI IHLCLIEBMPHSAIAKULPENENNETGLENVINKVATFKKPGVSGHYVE 822
 DB 773 AATDREIKREIHQLSIKPMHSELVKSLEPDKETGMESVIESVAHFKKPGLTGRMYE 832
 QY 823 LKDESLKDFNMYFYHYSKTQHSKASHMQKRRKQENKDEALPPPPPEPCFAPSKVINLL 882
 DB 833 LKPECAKEFNLYFYHFSRAEQSKAEQAQRKAKRENKEDTALPPALPPCFPLFASLVNLL 892
 QY 883 NCDIMWYILRTVFERAIDTDSNLWTEGMLQMAFHILAGLLBEKQOOLQKAPBEV--TPDF 941
 DB 893 QCDVMLYIMGTTLQWAVEHSGSAWSESMQLQVRLHLIGMALQEKHLENAVEGHVQTFP 952
 QY 942 YHKASRLGSSAMN---IQWLEKLAGIPQLEQCKDMITWLOMFDTVKELREKSLIVAT 998
 DB 953 TOKISKPGDAPHNSPSILAMLETQLNAPSLEAHKDMIRLLKQFNKAKIRE--CSSSP 1010
 QY 999 TSGSESIKNDIETHDKAERKRKAARLHRQKIMQASALQKNPIETHKLMYDNTSEM 1058
 DB 1011 VAEAGTMEESRRDKAERKRKAARLHRQKIMQASALQKNPIETHKLMYDNTSEM 1070
 QY 1059 PKGSEIMBEESTPAVSYSRIALQGRGPSVTEKEVLTFCILQBEQEVKIENNAMVLSA 1118
 DB 1071 DTSASATL--DSSPPVSDAALTALGPAQTVPEPRQVFTVTCILQBEQEVTVGSRAMVLA 1128
 QY 1119 CVQKSTALTQHRGKIELSGEALDPLFMDPDLAYGTYTSCGHHVMAHCWQRYFAVQLS 1178
 DB 1129 FVQRSTVLSKORTKI-ADPEKYDPLFMHPDLSGTHTSCGHHVMAHCWQRYFDSVQAK 1187

1179 SQ-----RIHWDLPDESGLVCPCKSLKNTVPIIPLOPQKINSENADALQLLTIA 1233
1188 EQRQORLRLHTS-YDVENGEFLCPCLBCLSNTVPLLL-LPPRSILSRRLN-FSQPDPUA 1244
1234 RWTOTLARISGYNI---RHAKGN-----PIPIPFNQMGDSITLFFHSILSF 1278
1245 QWTRATVQIKVQVMLRKHNAADTSSEDTAMNIIPIPEGFRP-----DFY----- 1292
1279 GVSEIKYSNISKEMVILFATTIYRIGLVKVPDPDRPVMLTWTSCATTIQAENLLGD 1338
1293 ---PRNPYSDSIKEMLTFTCTAAKYVGLKVHPNEGDRPVILCWGTCAVTIOGIERILSD 1349
1339 EGKPLEGALQNRHNGCLKALMQPAVAQRITCQVLIQKHLVRLSVLVNPKISEDTPCLL 1398
1350 BEKPFVFGPLFCRLDDCLSLTRPAAAHWTVALPVVQGHFCKLFASLVSDSDYDFPCLL 1409
1399 SIDPLHVLGAVLAPPSLYWDDPVDLQPSVSSVSNHLYLFLHTMAHMLQILL---TVD 1455
1410 DIDMFHLLVGLVAFALQCOD-----FGSSSLATGDLHIFHLVTWHAHVQILLTSCTEE 1464
1456 TGLPLAQVQDSEAHASFPABISQYTSIGSGDIPGWYLVWSLKNIGITPVLRCAALF 1515
1465 NGMD--QENPTGSEELAILSLHKTLLHQYTGSALEAPSGWHLWRSVRAAIMPLKCSALF 1522
1516 FHYLLGVTPEELHTNSABGEYSALCSYLSLPTNLLPQYWDVTRPILQRCADPALL 1575
1523 FHYLVNGVAPPDLQV-SGTSHFELHCLNYSLSPTNLHLFOENSIDNWSLIESQNSVEK 1581
1576 NCLKQKNTVVRYPKNSLIELPDDYSCLLNQASHFRCPSADDERKHPVLCFLCQAILC 1635
1582 RYLVNGERGAISYPRGANKLIDLPEDYSSLINQASNFCPSGSGDKSRAPTCLVCGSLC 1641
1636 SQNICQEIIVNGEVGACIFHALHCGAGVCFILKIRECRVVLVEGKARGCAYPAPYLDY 1695
1642 SQSYCCQAELEGEDVGACTAHTYSCGSGAGIPLRVRECQVFLAGTKGCFYSPYLDY 1701
1696 GETDPGLKRGNPLHLSRERYRKLHLVWOCHIEETARSOETNOMLPGFNWOLL 1749
1702 GETDQGLRGNPLHLCOERFRKIKLWQOHSITEEIGHAQEANOQLVGDIMOHL 1755

RESULT 7

Q8BN40 PRELIMINARY; PRT; 849 AA.
AC Q8BN40;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Mus musculus activated spleen cDNA, RIKEN full-length enriched library, clone:F830005C07 product:ubiquitin protein ligase E3 component n-recogin 1, full insert sequence. (fragment).
GN Names=Ubr1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Activated spleen;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning."
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Activated spleen;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection."
RN [3]
RP SEQUENCE FROM N.A.

RC STRAIN=NOD; TISSUE=Activated spleen;
RA The FANTOM Consortium,
RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Activated spleen;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M., Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Activated spleen;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P., Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M., Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K., Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M., Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J., Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD; TISSUE=Activated spleen;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P., Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W., Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T., Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kouda M., Koya S., Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M., Nishi K., Nomura K., Nunazaki R., Ohno M., Ohsato N., Okazaki Y., Saio R., Saichoh H., Sakai C., Sakai K., Sakazume N., Sano H., Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M., Tagawa A., Takahashi F., Takaku-Akaira S., Takeda Y., Tanaka T., Tomaru A., Tova T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK089616; BAC40933.1; -;
DR MGD; MGI:1277977; Ubr1.
DR GO; GO:0000151; C:ubiquitin ligase complex; IGI.
DR GO; GO:0005515; F:protein binding; IGI.
DR GO; GO:0004840; F:ubiquitin conjugating enzyme activity; IGI.
DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IGI.
DR GO; GO:000512; P:ubiquitin cycle; IGI.
DR GO; GO:0006511; P:ubiquitin-dependent protein catabolism; IGI.
DR InterPro; IPR000307; Ribosomal_S16.
DR Pfam; PF02207; zf-Ubr1; 1.
DR SMART; SM00396; Znf UBR1; 1.
DR PROSITE; PS00732; RIBOSOMAL_S16; UNKNOWN_1.
DR Ligase.
KW NON TER
FT NON TER
SQ SEQUENCE 849 AA; 98231 MW; B957FB7E5D08A89A CRC64;
Query Match 46.2%; Score 4262; DB 2; Length 849;
Best Local Similarity 94.3%; Pred. No. 1.1e-241;
Matches 801; Conservative 20; Mismatches 28; Indels 0; Gaps 0;
QY 1 MADEEAGGTERMEISAELEPQTPQRLASWWDQVDFYATFLHQAQLVPEIYFAEMDPDL 60
DB 1 MADEEMDGAERMDVSEPPPLAFQRPASWWDQVDFYATFLHQAQLVPEIYFAEMDPDL 60
QY 61 KQESVQMSIFTPLEWYLFEGEDPDICLKLKHSQFOLCGRVKSGTYSYSCDCAIDPT 120
DB 61 KQESVQMSILTPLEWYLFEGEDPDICLKLKHSQFOLCGRVKSGTYSYSCDCAIDPT 120

```
QY 121 CVLMDCFODSVHKNHRYKMHSTSTGGPCDGTAEWKTGFCVNVNHPGAGTICKNSRC 180
DB 121 CVLMDCFODSVHKNHRYKMHSTSTGGPCDGTAEWKTGFCVNVNHPGAGTICKNSRC 180
QY 181 PLNEEVIVQARRIFPSPVIVKVVEMTWESEKELPELQIREKNERYCYLVFNDEHSHYD 240
DB 181 PLNEEVIVQARRIFPSPVIVKVVEMTWESEKELPELQIREKNERYCYLVFNDEHSHYD 240
QY 241 VIYSQALDCELAELAQHTTADKGGRAVAKAGAAACORAKEDIKSHENVSQPLHV 300
DB 241 VIYSQALDCELAELAQHTTADKGGRAVAKAGAAACORAKEDIKSHENVSQPLHV 300
QY 301 EVLHSEIMAHOKFALRGSMWNKIMSYSDPFOIFCOACLRBPDSNPCLISRLMLWA 360
DB 301 EVLHSEIMAHOKFALRGSMWNKIMSYSDPFOIFCOACLRBPDSNPCLISRLMLWA 360
QY 361 KLYKGARKILHELIFSPFMEYKGLFAMEFVKYKQKQKQKQKQKQKQKQKQKQKQKQK 420
DB 361 KLYKGARKILHELIFSPFMEYKGLFAMEFVKYKQKQKQKQKQKQKQKQKQKQKQK 420
QY 421 MFTVPTLARHLIBEONVISVITETLLEVLPEYLDNRNKNFQGYSDQKLGRRVAVICDLK 480
DB 421 MFTVPTLARHLIBEONVISVITETLLEVLPEYLDNRNKNFQGYSDQKLGRRVAVICDLK 480
QY 481 YILISKPTITWELRMQFLEGESFLKILTCMQGMEIRRONCOHIEVDPDWEAAIAOM 540
DB 481 YILISKPTITWELRMQFLEGESFLKILTCMQGMEIRRONCOHIEVDPDWEAAIAOM 540
QY 541 QLNKILLMPEQWACDEBELLVAYKECHKAVMRCSTSFISSTKVVSQCHSLSTKSYRV 600
DB 541 QLNKILLMPEQWACDEBELLVAYKECHKAVMRCSTSFISSTKVVSQCHSLSTKSYRV 600
QY 601 SEDLSIHLPLSRTLAGLHVRLSRLGAVSLHFEVFPDQVEVLVEYPLRCLVLVAQVV 660
DB 601 SEDLSIHLPLSRTLAGLHVRLSRLGAVSLHFEVFPDQVEVLVEYPLRCLVLVAQVV 660
QY 661 AEWRRNGLSLSQVFFYQDVCKREEMDKDIIIMQIGASLMDPNKFLVLVQRYELAA 720
DB 661 AEWRRNGLSLSQVFFYQDVCKREEMDKDIIIMQIGASLMDPNKFLVLVQRYELAA 720
QY 721 FNKTSTKQDQILKQNTLIEEMQLVLYIVGRRYVPGVGNVTBKVMTREIHLICIEP 780
DB 721 FNKTSTKQDQILKQNTLIEEMQLVLYIVGRRYVPGVGNVTBKVMTREIHLICIEP 780
QY 781 MPHSATAKNLPENNETGLENVINKVATFKPGVSGHGYBLKOBLSKDFNMFFYHSK 840
DB 781 MPHSATAKNLPENNETGLENVINKVATFKPGVSGHGYBLKOBLSKDFNMFFYHSK 840
QY 841 TOHSAKHEM 849
DB 841 TOHSAKHEM 849
```

RESULT 8

```
ID Q6WKZ8 PRELIMINARY; PRT; 1756 AA.
AC Q6WKZ8;
DT 05-JUL-2004 (T-EMBLrel. 27, Created)
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
DE Ubiquitin ligase UBR2.
GN Name=Ubr2;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22948832; PubMed=14585983;
RX DOI=10.1128/MCB.23.22.8255-8271.2003;
RA Kwon Y.T., Xia Z., An J.Y., Tsaki T., Davydov I.V., Seo J.W.,
RA Sheng J., Xie Y., Varshavsky A.;
RT "Female lethality and apoptosis of spermatocytes in mice lacking the
```

RT Ubiquitin ligase of the N-end rule pathway.";

Mol. Cell. Biol. 23:8255-8271 (2003).

EMBL; AY280958; AAQ17202.1; -

GO; GO:0005515; F:protein binding; IPI.

DR InterPro; IPR002345; Lipocalin.

DR InterPro; IPR000408; Reg chr condens.

DR InterPro; IPR003126; Znf_Nrecogin.

DR Pfam; PF02207; zf-UBR1; 1.

DR SMART; SM00396; Znf_UBR1; 1.

DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_1.

DR PROSITE; PS00626; RCC1_2; UNKNOWN_1.

KW Ligase.

SQ SEQUENCE 1756 AA; 199392 MW; 646991E910776E18 CRC64;

Query Match 45.7%; Score 4218.5; DB 2; Length 1756;

Best Local Similarity 46.1%; Pred. No. 1.2e-238;

Matches 828; Conservative 343; Mismatches 539; Indels 85; Gaps 29;

QY 1 MADE---EAGGTER--MEISAEPLQTPORLASWDDQVDFYAPLHLAQLVPEIYPAEM 55

DB 1 MASEMEPEVOAIDRSLLCSAE-----BIAGRWLQATDLNREVYQHLAHCVPKICYCRGP 54

QY 56 DPDLKQBSVSQMSIPTPLEWYLFGBDPCDLEKLEKHSQ-AFOLCGRVFKSGTYYSCRD 114

DB 55 NPFPLEKEDTLAHLILGPNWTICADPALGPKLEQANKPSHLGCRVFAVGEPYSCRD 114

QY 115 CAIDPTCVLMDCFODSVHKNHRYKMHSTSTGGPCDGTAEWKTGFCVNVNHPGAGT 174

DB 115 CAVDPTCVLMDCFODSVHKNHRYKMHSTSTGGPCDGTAEWKTGFCVNVNHPGAGT 174

QY 175 -KENSRCPLNEEVIVQARRIFPSPVIVKVVEMTWESEKELPELQIREKNERYCYLVFN 233

DB 175 EEDPLVHLSEVIVARTINPAIFMRYAVDILTWEKSELDELVAESDYYCYMLFND 234

QY 234 EHSYDHYVYSIQRALDCELAELAQHTTADKGGRAVAKAGAAACORAKEDIKSHEN 293

DB 235 EVHTTEQVLYIQKAVNCKQKRAIGPATTVDRGRSVIRIGFYQYCDQAKTVLVRISQ 294

QY 294 SQHPLHVEVLHSEIMAHOKFALRGSMWNKIMSYSDPFOIFCOACLRBPDSNPCLIS 353

DB 295 TK-PLKQVQVHSSVAHQNFGLKALSWLGSVIGYDGLRRLICQVGLQSGPDGESSLD 353

QY 354 RLMLWDKLYKARKILHELIFSPFMEYKGLFAMEFVKYKQKQKQKQKQKQKQKQKQK 413

DB 354 RLMLWDKLYKARKILHELIFSPFMEYKGLFAMEFVKYKQKQKQKQKQKQKQKQKQK 413

QY 414 ITALSQVQVPTVTLARHL-IBEQNVISVITETLLEVLPEYLDNRNKNFQGYSDQK 469

DB 414 VADLSVQIPTVPSLAPNAPHRKKNLMTVILKAFMDHL-KHRDAQGRFQFERYTALQAF 472

QY 470 GRVYAVICDLKYLISKPTITWELRMQFLEGESFLKILTCMQGMEIRRONCOHIEV 529

DB 473 REVQSILDLKVLVLSKPTWSDDELROKFLQGFADFLKLLCMQMGNDPITRQVGHIE 532

QY 530 PDWEAAIAIQMOLKNILLMPEQWACDEBELLVAYKECHKAVMRCSTSFISSTKVVS 588

DB 533 PEWEAAFTLQMKLTHVISHVQVQWDCALDEKVLTEAVKCLAVLTQCHGGTDSQPITLS 592

QY 589 CGHSLSTKSYRVSEDLVSLHPLSRTLAGLHVRLSRLGAVSLHFEVFPDQVEVLVE 648

DB 593 CGHSETTRYCVSQBKVSILHPLSRLLAGLHVLLSKSEVAYKPPPELPLISELSPMLI 652

QY 649 PLRCLVLVAQVVAEMWRNGLSLISQVFFYQDVCKREEMDKDIIIMQIGASLMDPNK 708

DB 653 PLRCPVLCAQVHAGWRNGLSVLNQIYYHNKCREREMFDDKIVMLQTVGSMMDNHL 712

QY 709 LVLQRYELAAFN-----KTISTK--DODLIKQNTLIEEMQLVLYIVGRRYVPGVGN 761

DB 713 MIMLSRFBELYQLFSTPDYKGRPSSEVTHKDVVQNNLTIEEMLYLIIIMLVGERFNP 772

QY 762 VTKEEVWREIHLICIBPMHSAKILPENNETGLENVINKVATFKPGVSGHGYBLK 821

DB 773 VAATDEIKREIHLIQLSIKPMHSELVSLPDEKNTGMSVIESVAHFKKPGLTGRGM 832

QY 564 YKECHKAVMRCSTPSSSSKTVVQSCGHSLETKSYRVEDLSIHLPLSRTLAGLHVRLS 623
 DB 361 YKECHKAVMRCSTPSSSSKTVVQSCGHSLETKSYRVEDLSIHLPLSRTLAGLHVRLS 420
 QY 624 RLGAVALRLHEFVSFEDFQVEVLVEYPLRLCLVLVAQVVAEMWRRNGLSLSQVFFYQDVKC 683
 DB 421 RLGAVALRLHEFVSFEDFQVEVLVEYPLRLCLVLVAQVVAEMWRRNGLSLSQVFFYQDVKC 480
 QY 684 REBMYDKDIIMLQIGASLMDPNKFLLLVLRQYELAEAFNKTISTKDDLIKQNTLIBEM 743
 DB 481 REBMYDKDIIMLQIGASLMDPNKFLLLVLRQYELAEAFNKTISTKDDLIKQNTLIBEM 540
 QY 744 LQVLIIVGERRYVPGVNTKBEVTWREIHLICIEPMPHSAIAKULPENENNETGLENV 803
 DB 541 LQVLIIVGERRYVPGVNTKBEVTWREIHLICIEPMPHSAIAKULPENENNETGLENV 600
 QY 804 INKVAIFKFKPGVSGHGYELKDBSLKDFNNYFYHYSKTOHSAKEMKRRKQENKDEAL 863
 DB 601 INKVAIFKFKPGVSGHGYELKDBSLKDFNNYFYHYSKTOHSAKEMKRRKQENKDEAL 660
 QY 864 PPPPPPCFAPSKVINLLNCDIMMYILRTVPERAIDTDSNLWTEGMLQWAFHILALGLL 923
 DB 661 PPPPPPCFAPSKVINLLNCDIMMYILRTVPERAIDTDSNLWTEGMLQWAFHILALGLL 720
 QY 924 BEKQOLQKAPEREVTFDYHKASRLGSSAMNTOMLEKLGIPLEGQKDMITWILQMF 983
 DB 721 BEKQOLQKAPEREVTFDYHKASRLGSSAMNTOMLEKLGIPLEGQKDMITWILQMF 780
 QY 984 TVKRLREKSLIVATTSGSEIKNDEITHDK 1014
 DB 781 TVKRLREKSLIVATTSGSEIKNDEITHDK 811

RESULT 10

O15057
 ID O15057 PRELIMINARY; PRT; 1275 AA.
 AC O15057;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE KIAA0349 protein (Fragment).
 GN Name=KIAA0349;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=97349984; PubMed=9205841;
 RA Nagase T., Ishikawa K., Nakajima D., Ohira M., Seki N., Miyajima N.,
 Tanaka A., Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. VII.
 The complete sequences of 100 new cDNA clones from brain which can
 code for large proteins in vitro."
 RL DNA Res. 4:141-150(1997).
 DR EMBL; AB002347; BAA20806.1;
 DR GO; GO:0000151; C:ubiquitin ligase complex; IEA.
 DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
 DR GO; GO:0008270; F:zinc ion binding; IEA.
 DR GO; GO:0016567; P:protein ubiquitination; IEA.
 DR InterPro; IPR000408; Reg chr condens.
 DR InterPro; IPR001841; Znf_ring.
 DR SMART; SM00184; RING_1.
 DR PROSITE; PS00626; RCCL_2; UNKNOWN_1.
 FT NON_TER 1
 SQ SEQUENCE 1275 AA; 145331 MW; 101FF1F66E8056066 CRC64;

Query Match

34.3%; Score 3160; DB 2; Length 1275;

Best Local Similarity 47.3%; Pred. No. 1.2e-176;

Matches 612; Conservative 248; Mismatches 393; Indels 40; Gaps 19;

QY 479 LKYLISKPTIWTERTLRLMQPLRSGPRSPKILTCMOGMEIRROVGOHIEVDPDWEAAIAI 538
 DB 1 LKYLISKPTIWTERTLRLMQPLRSGPRSPKILTCMOGMEIRROVGOHIEVDPDWEAAIAI 60
 QY 539 QMQLKNILMFQEWACDEBELLIVAYKECHKAVMRCSTPSSSKTVVQVOS-CGHSLETKS 597
 DB 61 QMQLKTHVISMMQDWACASDEKVLIEAYKCLAVLMQCHGGYTDGEQPTILSICGHSVETIR 120
 QY 598 YRVSEDLVSIHLPLSRTLAGLHVRLSRLGAVSRHLHFVSFEDFQVVLVEYPLRLCLVLVA 657
 DB 121 YCVSQBKSVIHLPVSRLLAGLHVLLSKSVAVKFPPELLPLSELSPMLIEHPRLCLVLCA 180
 QY 658 QVVAEMWRRNGLSLSQVFFYQDVVCKREEMVQKDIIMLQIGASLMDPNKFLLLVLRQYEL 717
 DB 181 QVHAGWRRNGPSLVNQIIYHNKCRREMPKDVVMLQTVGVSMDPNHFLMILSRPEL 240
 QY 718 ABAFN-----KTIATK--DQDLIKOYNTLIBEMQLVLIIVGERRYVPGVNTKBEVTW 770
 DB 241 YQIFSTPDYGRKPSSEITHKDVVQNNTLIEMLYLIIMLVGERFSPGVQVWATDEIKR 300
 QY 771 EIIHLCTIEPMPHSAIAKULPENENNETGLENVINKVATFKKPGVSGHGYELKDBSLKD 830
 DB 301 EIIHQLSIPMAHSELVKSLEPDKENKTCMBSVIEAVAHFKKPGLTGRGMYELKPECAKE 360
 QY 831 FNNYFYHYSKTOHSAKEMKRRKQENKDEALPPPPPCFAPSKVINLLNCDIMMYI 890
 DB 361 FNLIFYHFSRASQABEAQRKLKQNRDTPALPPVLPFPFCLFASLNIILQSDVMLCI 420
 QY 891 LRTVPERAIDTDSNLWTEGMLQWAFHILALGLLEKQOLQKAPERE-VTFDFVHKASRLG 949
 DB 421 MGTILQWAVEHNGYAWSESMRLQRLHILQMALQEKQHLNTEHVVTFVTQKLSKPG 480
 QY 950 SSAMN---IQMLEKLKGIPOLEGQKDMITWILQMPDTVKRLREKS-CLIVATTSGSEI 1005
 DB 481 EAPKNSPSILAMLETQLQAPYLEVHKDMIRWLKTFNAVKMRESSPTSPVATECTIM- 539
 QY 1006 KNDEITHKKAERKKAABARLHQKIMQAQNSALQKNIETHKLMYNTISEMPGKEDSI 1065
 DB 540 --EBSRDQKKAERKKAABARLHQKIMQAQNSALQKNIETHKLMYNTISEMPGKEDSI 597
 QY 1066 MEESBSTPAVDYSRIALGPKRSPVTEKEVLTCILCOBEQEVKIENNAMVLSACVQKSTA 1125
 DB 598 LDH--SPVNSDMTLTALGPTQVPEQRFVTCILCOBEQEVKESRAVLAFAVQSTV 655
 QY 1126 LTHQKPTIELSGEALDPLFMDPLAYGTYTSCGHHVAVCWQKYFEAVQLSSQQ---- 1181
 DB 656 LSKNRSKPIQ-DPEKYDPLFMDPLSCGTHHTSSCGHIMHACWQRYFDSVQAKEQRQR 714
 QY 1182 -RIHVDLFDLESGLVCLCKSLCNTVPIIIPLOPKINSENADALQLLTARVQTVL 1240
 DB 715 LRLHTS-YDVENGFLCPLCECLSNVPIPL-LPPRNIFNNRLN-FSDQPNLTQWIRTIS 771
 QY 1241 ARISGVNIRHAKGENPIPIFFNQMGDSTLPHSILSPGVSSIKYNSIKEMVILFATT 1300
 DB 772 QQIKALQFLRKEESTPNNASTKNSENVDELQIPEGRFPDRPKIPIYSESIKEMLTFFGA 831
 QY 1301 IYRIGLKVPPDERPRVPMLTWSTCAFTIATENLLGDEKPLFGALQWROHNGLKALMQ 1360
 DB 832 TVYKGLKVHPNEEDPRVIMCWGSCAYTIQSIERILSDEDEKPLFGPLCLDCLSLRTR 891
 QY 1361 FAVAORITCPQVLIQKHLVRLSVVLPNIKSBDTPCLLSIDILFHVILGAVLAPPSLYWD 1420
 DB 892 FAAAHTVAVSVVQGHFCKLFAVLVNDSEBELPCILIDMFHLLVGLVLAFFALQCOD 951
 QY 1421 PVDLQPSVSSSYNHLPLHILTMALMLQILL--TVDTGLPLAQVQEDSEAHSSSFF 1477
 DB 952 -----FSGISLGTGLHIFHLVMAHIIQILLTSTCEENGMD--QENPPCESESAVL 1004
 QY 1478 AEISOTSGSIGCDIP-GWILWVSLKNGITPYLRCAALFFHYLLGVTPPELHTNSAEGE 1536
 DB 1005 KTLHQYT-GSALKEIPSGWHLWRSYRAGIMPLPKCSALFFHYLVNGVSPSPDQV-PGTS 1062
 QY 1537 YSALCSYLSLPTNPLFLFQYWDVTVRLLQRCADPALLNCLUKQKNTVVYRPRKNSLIE 1596

```

Db      1063 FEHLCSYSLPNNLCLFOENSEINWSLIESCRNSEVRYLEGEDAIRYPRESKLN 1122
Qy      1597 LPDDYSCLLNQASHFRCPASADDERKHPVLCIFCGAILCSQNICQEIYNGEVCACIFH 1656
Db      1123 LPEDYSSLLNQASHFRCPASADDERKHPVLCIFCGAILCSQNICQEIYNGEVCACIFH 1182
Qy      1657 ALHCGAGVCIPIKIRECRVVLVEGKARGCAYPAPYLDEYETDGLKRGNPLHLRERYR 1716
Db      1183 TYSCSGVGIFLRVRECVLFLAGTKGCFYPPYLDYDGETDQGLRGNPLHLCKERPK 1242
Qy      1717 KLHLVWQOHCIIIEETARSQETNQLFGFNWQLL 1749
Db      1243 KIQLKWHQHSVTEIEIGHAQBANTLVGIDWQHL 1275

RESULT 11
ID      Q80U31 PRELIMINARY; PRT; 1246 AA.
AC      Q80U31;
DT      01-JUN-2003 (Tremblrel. 24, Created)
DT      01-MAR-2004 (Tremblrel. 26, Last sequence update)
DE      MKIAA0349 protein (Fragment).
GN      Name=MKIAA0349;
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]_TaxID=10090;
RP      SEQUENCE FROM N.A.
RC      TISSUE=Brain;
RX      MEDLINE=22579291; PubMed=12693553;
RA      Okazaki N., Kikuno R., Ohara R., Inamoto S., Aizawa H., Yuasa S.,
RA      Nakajima D., Nagase T., Ohara O., Koga H.;
RT      "Prediction of the coding sequences of mouse homologues of KIAA genes:
RT      II. The complete nucleotide sequences of 400 mouse KIAA-homologous
RT      cDNAs identified by screening of terminal sequences of cDNA clones
RT      randomly sampled from size-fractionated libraries.";
RL      DNA Res. 10:35-48(2003).
DR      EMBL; AK122254; BAC65536.3; -
DR      InterPro; IPR000408; Reg chr condens.
DR      PROSITE; PS00626; RC01_2; UNKNOWN_1.
SQ      NON TER
SQ      SEQUENCE 1246 AA; 140966 MW; C7A0ED7B2D9D265C CRC64;

Query Match 32.7%; Score 3014.5; DB 2; Length 1246;
Best Local Similarity 45.9%; Pred. No. 4e-168;
Matches 599; Conservative 238; Mismatches 374; Indels 95; Gaps 22;

Qy      479 LKYLISKPTIWTIRLMQFLEGFRSLKILTCMQGMBEIRROVQGHIEVDPDWEAAIAI 538
Db      1 LKVLISKPTWESDELRLQFQGFDAFLKLCMQGMDPIRQVQGHIEPEWEAAFTL 60
Qy      539 QMQLKNILMFQWACDEBELLVAVKECHAVRSTSFISSTKVQOS-CGHSLETKS 597
Db      61 QMQLKTHVISMVDQWACALDEKVLIEAYKCLAVLTQCHGGTDCGEQPTILSICGHSVETIR 120
Qy      598 YRVSEDLVSLPLSRKTLAGLRVLRGLGAVSRILHEFVSFEDQVEVLVEYPLRCLVLVA 657
Db      121 YCVSEKVSILHPLSRLLAGLHVLKSEVAYKFPBELLPLSELSPMLIEHPLRCLVLCA 180
Qy      658 QVVAEMWRNGLSLISQVFYQDYQCKREEMDYKDIIMLQIGASLMDPNKPLLLVLRQYEL 717
Db      181 QVHAGMWRNGLSPVSLVQIYYVHNKCRREMFDDKDIIVMLQTVGSMDPNHPLMTLSRFEL 240
Qy      718 AEAFN-----KTIETK-----DQDLIKQNTLIEEMQLVLIYVGERYPGVGNVKEEYVTR 770
Db      241 YQLFSTPDYGRFSSEVTHKDVVQNNLTIEEMLYLIMLVGERFNPFGVQVAATDEIKR 300
Qy      771 EITHLICIEPMPSAIAKNLPENNETGLENVINKVATPKPGVSGHYVELKDESLKD 830
Db      301 EIIHQLSIKPMASSELVKSIPEDB-----ELKPCAKE 333

```

```

Qy      831 FNMVYHYSKTQHSKAEHQKRRKQENKDEALPPPPPEFCPAFSKVINLLNCDIMMYI 890
Db      334 FNLFIYHSRAQKAEQAQRKLRKREKEDTALPPALPPFCPLFASLVNLIQCDVMIYI 393
Qy      891 LRTVFERAIDTSDNLWTEGMLQMAFHIALGLLBEKQOLQKAPBEV-TFPDYHKAIRUG 949
Db      394 MGTILQWAVEHHGSAWSSMLQVLLHGLMALQBEKHLNENAVEGHVQVTFPTQKISKPG 453
Qy      950 SSAMN---IOMLEKIKGIPOLEGKQMIWILQFQDMITWILQFQDMITWILQFQDMITW 1006
Db      454 DAPHNSPILAMLETQNPASLEAKHDMIRLLKMFNAIKIRE--CSSSSPVAEAGTI 511
Qy      1007 NDEITHDKEKAEKKEKAEARLRKQIMQAQNSALQKQFIETHKLMYDNTSEMPGKEDSIM 1066
Db      512 MEESRDQKAEKKEKAEARLRKQIMQAQNSALQKQFIETHKLMYDNTSEMPGKEDSIM 571
Qy      1067 BEESTPAVDYSRIALGPKRGPSVTEKEVLTICLQBEQEVKIKENNAMVLSACVQKSTAL 1126
Db      572 --DSSPPVSDAALTALGFTQVPE--PRQFVTCILQBEQEVTVGSRAMVLAAPVQRSTVL 627
Qy      1127 TQHRKPIELSGEALDPLFMDPDLAYGTYTSCGHVMAVCHWKYFEAVQLSSQ----- 1181
Db      628 SKDRTKTI-ADPEKYDPLFMHPDLSCGTHTSCGHVMAHCHQRYFDSVQAEQRQORL 686
Qy      1182 RIHVDLFLDSGEYLCPCKSLCNTVPIIPILOPOKINSADALQALLTLARWITQVLA 1241
Db      687 RUHTG-YDVENGFEFLCPCBCLSNVIFEL-LPPRSILSRRLN-FSDQPDLAQWTRAVTQ 743
Qy      1242 RISGYNI---RHAKGEN-----PIPIFNQMGDSTLPHSILSGVSESSIKY 1286
Db      744 QIKVQMLRRKHNAADTSSEDETEAMNIIPIEGFRP-----DFY-----PRNPY 788
Qy      1287 SNSIKEMWILFATTIYRIGLKVPPERPRVPMLTWSTCAPTIOAIENLGLDEGPLECA 1346
Db      789 SDSIKEMLTTFGTAAAYKVLKVPNEGDPVILCMGTCAVYTIQIERILSDEERKPVFP 848
Qy      1347 LQNRQHNGKALMQFAVQRTCPQVLIQKHLVLLSVLPNPKIKNEDTPCLLSIDLPHVL 1406
Db      849 LPCRDDCLRLSUTRFAAAHWTVALLPVQGHFKLPASLVPSSDSYEDLPCILIDIMFHL 908
Qy      1407 VGAVLAFPSLYWDDPVDLPQSSVSSYNHLYLFLHITWAHMLQILL---TVDTGLPLAQV 1463
Db      909 VGLVLAFFALQCOD-----FSGSSLATGDLHIFLHVTMAHIVQIILLTSCTEENGMD--QE 961
Qy      1464 QDSSEASASFPFAEISQYTSIGCDIPGHVYLVSLKNGITPYLRCAALFHYLLGYT 1523
Db      962 NPTGEEELAILSLHKLTHQYTGSAKLEAPSGHMLWRSVRAAIMPFLKCSALFHYLNGVP 1021
Qy      1524 PPEELHTNSAEGEYSALCSYLSLPTNLPLLPQYVDTVRPILQRCWADPALLNCLKQKNT 1583
Db      1022 APDDIQV-SGTHFHELCNLSLPTNLHLFQNSDIMNSLIESWCQNSKRYLNGERG 1080
Qy      1584 VVYPRKRNLSLIEPDDYSCLLNQASHFRCPASADDERKHPVLCIFCGAILCSQNICQOE 1643
Db      1081 AISYPRGANKLIDLPEYSSLLNQASHFRCPASADDERKHPVLCIFCGAILCSQNICQOE 1140
Qy      1644 YNGEVGACIFHALHCGAGVCIPIKIRECRVVLVEGKARGCAYPAPYLDEYETDGLK 1703
Db      1141 ELEGEDVGACTAHTYTSYSCSGAGIFLRVRECVLFLAGTKGCFYPPYLDYDGETDQGLR 1200
Qy      1704 RGNPLHLRERYRKLHLVWQOHCIIIEETARSQETNQLFGFNWQLL 1749
Db      1201 RGNPLHLRERYRKLHLVWQOHCIIIEETARSQETNQLFGFNWQLL 1246

RESULT 12
ID      Q8K216 PRELIMINARY; PRT; 1109 AA.
AC      Q8K216;
DT      01-OCT-2002 (Tremblrel. 22, Created)
DT      01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT      01-MAR-2004 (Tremblrel. 26, Last annotation update)

```

DE E130209G04Bik protein.
GN Name=E130209G04Bik;
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RX NCBI_TaxID=10090;
ON [1]
RN
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Cassavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skaleka U., Smalusz D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RW [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Mammary tumor. C3;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC031403; AAH31403.1; -
DR MGD; MGI:1925978; E130209G04Bik.
DR GO; GO:0005515; F:protein binding; IPI.
DR InterPro; IPR000408; Reg chr condens.
DR InterPro; IPR001841; Znf_ring.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00626; RCC1.2; UNKNOWN 1.
SQ SEQUENCE 1109 AA; 125288 MW; B506F24067283279 CRC64;

Query Match 29.1%; Score 2687.5; DB 2; Length 1109;
Best Local Similarity 46.8%; Pred. No. 5.5e-149;
Matches 533; Conservative 209; Mismatches 333; Indels 65; Gaps 19;

QY 644 VLVEYPLRLCLVLAQVVAEWRNGLSLISQVYQDVCKREBYDKDIIMLQIGASLMD 703
DB 1 MLIEHPLRLCLVLAQVVAEWRNGLSLISQVYQDVCKREBYDKDIIMLQIGASLMD 60

QY 704 PNKFLLLVLRVELAEAFN-----KTIATK--DODLTKOYNTLIEMLQVLYVIGERYV 756
DB 61 PNHFLMILSRFELYQLFSTPDYGRKFSSEVTHKVVQQNNTLIEMLYLIIMLVGERFN 120

QY 757 PGVGNVTKSEVTRREIHLICIEPMPHSAIAKNLPENNNETGLENNVINKVATFKKPGVS 816
DB 121 PGVQVAANDIEKREIHLQSLKPMASSELVSLPEDEKNETGSMESVIESVAHFKKPGIT 180

QY 817 GHGVTELKDESUKDFNMYFYHYSKTOHSAKHMQRKQKQKQKQKQKQKQKQKQKQKQK 876
DB 181 GRGVTELKPECAKEFNLYFYHFSRAEQSKAEBAQRKLEKEDTALPPFPALPPFCPLPA 240

QY 877 KVINLNCIDIMVILATVERAIDTNSLWTEGMLQMAFHILALGLLEEKQLOKAPREE 936
DB 241 SLVNTILOQDMVIMYIMTILQWAVEHNGSAWSESLQRLVHLIGMALQSEKHLNNAVEGH 300

QY 937 V-TFDPYHKASRLGSGAMN---IQMLEKLKGIPOLEGOKDMITWTLOMFDVTKRLREKS 992
DB 301 VOTFTFTQIKSPGDAPHNSPILAMLETLOVAPSLEAKDMIRWLLKMFNAIKKIRE-- 358

QY 993 CLIVATTSGSESINKDEITHDKEAERKKAERKKAERKKAERKKAERKKAERKKAERKKA 1052

RESULT 13

Q7QEV6 PRELIMINARY; PRT; 1679 AA.
ID Q7QEV6
AC Q7QEV6;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Eb1F8105 (Fragment).
GN Name=eb1F8105; ORPNames=ENSANG00000006115;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (Mar-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.

Db 1650 -QOPDARRSLQVVPCLRLPLRLKVLCDPFDLSLNSVSDIFCPNNEEREEMKTPMCLICGL 1708
 Qy 1633 ILCSQNICCOEIVNGEVGACIFALHCGAGVCFILKIBECRWLVGKARGCAYPAPYL 1692
 Db 1709 ILGQSYCCQPELGRKSVGACHTHAHACGAEGVIFLIRIDCQVYVL-GRGKCGCFVPPPYL 1767
 Qy 1693 DEYGETDPLGRGNPLHLGRYRKLHLVQOHCIEETARSQETQMQLFGFNW 1746
 Db 1768 DEYGETDMGRGNPLRLSQAAYRKIYLQWLHGLHGEIARLNDANVAQAQW 1821
 RESULT 15
 Q9YK91
 ID Q9YK91 PRELIMINARY; PRT; 1824 AA.
 AC Q9YK91;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE CG9086-PA.
 GN ORFNames=CG9086;
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 RN [1]
 SEQUENCE FROM N.A.
 RP MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H., Blazey R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
 RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolach R.S.,
 RA Borkova D., Botchan M.R., Bouck J.J., Brokstein P., Brotier P.,
 RA Burtis K.C., Busam D.A., Butler H., Davenport L.B., Davies P.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foslberg C., Gabrieli A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
 RA Jalali M., Kalush P., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D., Scheeler P., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirkas R., Tector C., Turner C., Venter E., Wang A.H., Wang X.,
 RA Wang Z.Y., Wasmann D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodgett W., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhou X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195(2000).
 RN [2]
 SEQUENCE FROM N.A.
 RP MEDLINE=22426065; PubMed=12537568;
 RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
 RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
 RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,

RA Svirkas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
 RA Weinstock G., Scher S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
 RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila
 melanogaster euchromatic genome sequence.";
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
 RN [3]
 SEQUENCE FROM N.A.
 RP MEDLINE=22426070; PubMed=12537573;
 RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirkas R.,
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
 RA Ashburner M., Celniker S.E.;
 RT "The transposable elements of the Drosophila melanogaster euchromatin:
 a genomics perspective.";
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
 RN [4]
 SEQUENCE FROM N.A.
 RP MEDLINE=22426069; PubMed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a
 systematic review.";
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
 RN [5]
 SEQUENCE FROM N.A.
 RP FlyBase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RN [6]
 SEQUENCE FROM N.A.
 RP FlyBase;
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF003504; AAF48687.2; -.
 DR FlyBase; FBgn0030809; CG9086.
 DR GO; GO:0000151; C:ubiquitin ligase complex; IEA.
 DR GO; GO:0004842; F:ubiquitin-protein ligase activity; IEA.
 DR GO; GO:0008270; F:zinc ion binding; IEA.
 DR GO; GO:0016567; P:protein ubiquitination; IEA.
 DR InterPro; IPR003126; Znf Nrecoqin.
 DR InterPro; IPR001841; Znf ring.
 DR Pfam; PF02207; zf-UBR1; I.
 DR SMART; SM00184; RING_1.
 DR SMART; SM00396; ZNF_UBR1; 1.
 SQ SEQUENCE 1824 AA; 208359 MW; 79992C11175E82A0 CRC64;
 Query Match 24.7%; Score 2276.5; DB 2; Length 1824;
 Best Local Similarity 31.7%; Pred. No. 1.6e-124;
 Matches 607; Conservative 307; Mismatches 707; Indels 293; Gaps 56;
 Qy 5 EAGGTERMEISAEPLPQTPORLASWWDQVDFYTAFLHLAQLVPEIYFAEMDPDLKQEE 64
 Db 29 QAGTLDRSDIIEFLKRESPK---YFDYQT---SATVKDTNVTILKCMPEK---SLAKEE- 78
 Qy 65 SVQMSIFTPLEWLVGEPDIDICLKLKHSQ-AFOLCGRVKSGSETTSCDCAIDPTCVL 123
 Db 79 -----IIDVVVFMLGNDNPSLLEKRLLEGNTATVCGKVKNGEPTYSRCGVDPTCVL 133
 Qy 124 CMDCFDSDVHKHRYKMHSTGCGFCDCGDTBAWTKGPPCVNHPGPRAGTITKENSRCPLN 183
 Db 134 CYNCFKSAHRFKYKMSISGGGCCDCGDDAANKQOYCELHLANKNPLESKI---LT 190
 Qy 184 BEVIVQARKIPPSVIKYVVMTEWBEKEKLPPELQIREKN-----BRYCVLPNDEH 235
 Db 191 DAVLERVEICFCAILAFCSYVL-----EIEPNASLQCLDGNVGGQVDGAQYCTVLYNDES 246
 Qy 236 HSYDHVYISLQALDCELAELTAIDKEGRVAKGAYAAQCEAKEDIKSH----- 289
 Db 247 HTFDQVIQTLTKIAKRAKADMEIVAAIDREGRAVVKCTDFEENCKLKVSIENMLPTS 306
 Qy 290 --SENVSQHPHLHVEVLHSEIMAHQKFAIRLGLSGWMNKIMSYSDFRQIFCOACLRBEPDSE 347

Db 307 LVSTARNQSLRTSVLHIGVACQFALQQLGFWQFVVRHYLFRKTFSELVORKQ---E 363
Qy 348 NPCLISRLMDWLAKYKARKILHELIFSPFMEYKCLFAMEFVKYKQLOKEVISDD 407
Db 364 TFC-IRHILEYDVKLAKTARTCWHRLLISGLMEYDNKMLAQOFSRRYATIVEDFISDD 422
Qy 408 HDRSISITALSVQMFVPTLRLHILBEQNVISVITETLLEVLPEYLDNRNKNFQO---YS 465
Db 423 HDHAFSIVLSVLQFTVPSTAHLLIAHEGIFDKLLHTFVHVAIEKTRNKLHFSKNIAS 482
Qy 466 QDKGLRYYAVICDKYLILSKPTIWTLELMQFLEGRSPFLKILTCMQGMEETRRQVGOH 525
Db 483 LTFKRNANYLYDLRYLLSLKUPVLSNDLNGLEGCRALMRVLNVQGNESRTOGTQH 542
Qy 526 IEYDPPWEAIAIOMQLAKNILLFQEWACDBELLVAYKECHKAVMKRSTSPSSKTV 585
Db 543 MDYEPWECAFNHLHKLATTISQVIDWASGDVLLKRLKYMTWALV--SNSFIUGKEV 600
Qy 586 VQS---CGHSLETYSRVSEDLVSIHPLSRTLAGLHVLRLGAVSRHLHFVSFEDPQV 642
Db 601 MOPKQVADHVANCLVDIVSQVPSIHLPLSRFPAGIYL---HLGA---HD-LTYDGLQT 652
Qy 643 EV-----LVEYPLRCLVLVAQVWAEWRNGLSLISQVFFYQDVKCEEMVYDKD 691
Db 653 ETALSIKLTPREIIBPVLTQAMIAQVAGLWRRNGYTLHLQLYFYRNVRCRVEMLDRD 712
Qy 692 IIMLOIGASIMDPNKLLLVLORYE---LAAEFNKTISTKO---QDLIKQYNTLIBEMIQ 745
Db 713 IACLOIGASLMESNEFLHVLNRPNTIPMLQENYWSLLSGNEMNDIIEB-ASIDPEFLE 771
Qy 746 VLIYVGERYVPGVGNVTKBEVTMRIIHLCLTEPMPSHAIAKNLPENE--NNETGLENV 803
Db 772 LLIVIIIGERWMPGVSMVTBEDRLKEIILLCIKPYSHSLSRALPDGNSGNSDNVFEV 831
Qy 804 INKAVTFKPP-GVSGHGVVELKDESLEKDNMYFYHYSKQHSKAEHMQKRRKQENKDEA 862
Db 832 INTVAVFKPPVGADSGVYELKEHLKBFNMYFYHTYKEDKSKABELQERRKAKGQOLVC 891
Qy 863 LPPPPPPPCFAPSKVINLINCDDIMMYILRTVPERAIDTDSNLWTEGMLQMAFHIALGL 922
Db 892 CPPPMLPLTPATPMPANILQCPVPLNICSLIMERALNAYSRSFTESHQVQLHLGYAI 951
Qy 923 LEEKQOLQKAPSEVTFDFYHKASRLGSSAMNITOMLEKLGIPQLEGQKDMITWLOMP 982
Db 952 QEE-----LSEHYPLSPFYERSQBYG-----ILEKLEBLARCPLEAHYDFVLWTIERF 1000
Qy 983 DTVKRLREK-----SCLIVATSGSESINKDEITHDKKAEKRRKAEARLHROK 1032
Db 1001 ---KOLQAKQAPSGDGRAGPSC-SQOGTGGKLSLSABE---QAREERENRARLAERRAH 1052
Qy 1033 IMAQMSALQKNPLETEHKLMDNTSEMPGK-----DSIMEBESTPAVDYSRIA-L 1082
Db 1053 IMAQMQAKQKFLISSNAEMFADTENETRESASTGPMWEDIPPEEQGAVALSKVACL 1112
Qy 1083 GPKRGPSTVEKSVLTICLQEOEVEKIERNNAMVLSACVOKSTALTQHRKPIELSGEALD 1142
Db 1113 GPDRKPYHGTDTDFKILCFENCAISRGGRQLVSSAFVQTSRV----- 1155
Qy 1143 PLFMDPDL---AYGTTGSGHVMHVCWKQYF---EAVQLSSQRIHVLDLFDLESGEYL 1196
Db 1156 -IFTTENLRSQSALHISCGGHVMYHYSWLEVFTEEPKELRPHENRAALAAQANVFQ 1214
Qy 1197 CPLCKSLCNTVPIPILOPQ-----KINGENADALAQL---LTLAR 1234
Db 1215 CPYCRTLSNAILPVTETLPAFSAAPPSPNESYPLDSFVIMSTLALTELGNVKDHETTLTP 1274
Qy 1235 WIQTVLARISGNIRHAKGENPIPIPFNOQMGDSTLEPHSILSFVGVSSIKYSNSIKENV 1294
Db 1275 SVSNIL-RLSGVVGGLAQFERSVOLIKNPP-----RLHADYEGIE-----PLKKA 1319
Qy 1295 ILPATTIYRIGLK-VPPDRDRPRVPMLTWSTCAFTIOAIENTLGDGKPLFGALONROHN 1353

Db 1320 LLNTMKIQOOSHLKDHPAIESIENVPIL-WDSCSYTLOALEIYIYAVEKPLKARLSMRHQS 1378
Qy 1354 GLKALMQ-----FAVAQRITCPQVLIQKHLVRLLSVVLNPKISEDTFCLLSIDLPHVLV- 1407
Db 1379 CARNLVRACSRSSALEWETDLPLPPMRQABFSSRLLDITIFNQNDTSVLEWDCFRLVLP 1438
Qy 1408 ---GAV-LAFPSLYDDPVDLQPSVSSSYNNHLYLFLHITMAHMLQILLTVDTGFLPLAQ- 1462
Db 1439 FQFGVLNLMVPEKGYKTIIP-----SGSMFDPYIMQTMFLAQLTRAVLCFDVEKEKAKR 1492
Qy 1463 -----VOEDSEER---AHS 1472
Db 1493 AEKAPNSELTDQDYIBQLFSRIRDNMDIFRRYNIIPARVLQKTKQKQOLVEESEENQGHG 1552
Qy 1473 -----ASSFPABISQYTSIGCDIPGWYLMVSLKNGITPYLRCAALFFHYLLGVTPP 1525
Db 1553 QTVVIPCESHHLALLLEY-----VQRQWSSFLRCSCLFYRFLTDVDP 1595
Qy 1526 BELHTNSAGEYSALCSYLSLPTNLELLEQVWDTV-RPLQWACADPALLNCLKQKNTV 1584
Db 1596 DTFTDQPD-RFDLMCQYLGLDPLMGVYFD--METVYATMHSFASHPHIDREVEQR--- 1649
Qy 1585 VRYPRKNSL-IE-----LPDDYSCILLNQASHFRCPRSADDERKHPVLCPCGA 1632
Db 1650 -COPDARRSLQVEPCURPLRKLKLCDDFSDLSNSVSDIFCPNNEREEMKPTMCLICGL 1708
Qy 1633 ILCSONICCOBIWNGEEVGCACIFHALHCGAGVCIFLKIRECRVVLVEGKARGCAYPAPYL 1692
Db 1709 ILCCQSYCCQPELKGKVSVCCTHAAHACGAEVGFILIRDCQVVYL-GRGKGFVPPPYL 1767
Qy 1693 DRYGETDPLGKGNPLHLSRERYRKLHLVWQOHCIIIEETARSOETNQMLFGFNW 1746
Db 1768 DRYGETDMGLRRGNPLRLSQAAAYRKTYLQWLGHGLHGETARLNDNANVAAAQW 1821

Search completed: July 9, 2005, 15:31:41
Job time : 168 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2005, 15:31:48 ; Search time 133 Seconds
(without alignments)
5080.389 Million cell updates/sec

Title: US-10-758-672A-2

Perfect score: 9224

Sequence: 1 MADEAGGTERMEISAEIPQ.....ETARSGETNQLGFWQLL 1749

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1726216 seqs, 386330316 residues

Total number of hits satisfying chosen parameters: 1726216

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US10F_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
21: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
22: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9224	100.0	1749	16 US-10-758-672A-2	Sequence 2, Appli
2	9224	100.0	1749	16 US-10-357-819-10	Sequence 10, Appl
3	9224	100.0	1749	17 US-10-758-636A-2	Sequence 2, Appli
4	9168	99.4	1738	14 US-10-287-218-1	Sequence 1, Appli
5	9168	99.4	1738	16 US-10-474-291-1	Sequence 1, Appli
6	9087	98.5	1734	16 US-10-758-672A-19	Sequence 19, Appl
7	9087	98.5	1734	16 US-10-758-636A-19	Sequence 19, Appl
8	8617	93.4	1757	16 US-10-758-672A-15	Sequence 15, Appl
9	8617	93.4	1757	17 US-10-758-636A-15	Sequence 15, Appl
10	4327	46.9	1755	16 US-10-758-672A-4	Sequence 4, Appli
11	4327	46.9	1755	17 US-10-758-636A-4	Sequence 4, Appli

12	4278	46.4	1755	16 US-10-758-672A-6	Sequence 6, Appli
13	4278	46.4	1755	17 US-10-758-636A-6	Sequence 6, Appli
14	3160	34.3	1275	16 US-10-755-889-635	Sequence 635, App
15	3066	33.2	594	15 US-10-094-749-2555	Sequence 2555, Ap
16	2710.5	29.4	1109	9 US-09-529-063-55	Sequence 55, Appl
17	2710.5	29.4	1109	14 US-10-414-378-55	Sequence 55, Appl
18	1169.5	12.7	452	15 US-10-264-237-2544	Sequence 2544, Ap
19	550	6.0	1270	16 US-10-437-963-193003	Sequence 193003,
20	468	5.1	147	15 US-10-140-047-2153	Sequence 2153, Ap
21	453.5	4.9	951	17 US-10-480-988-17	Sequence 17, Appl
22	414.5	4.5	203	15 US-10-264-237-1513	Sequence 1513, Ap
23	386.5	4.2	889	16 US-10-425-115-307970	Sequence 307970,
24	396	3.2	250	15 US-10-296-115-1087	Sequence 1087, Ap
25	295	3.2	247	14 US-10-050-704-1107	Sequence 110, App
26	295	3.2	247	16 US-10-798-512-110	Sequence 110, App
27	276	3.0	81	14 US-10-016-349A-227	Sequence 227, App
28	275	3.0	499	15 US-10-425-114-68139	Sequence 68139, A
29	259	2.8	550	15 US-10-424-599-182689	Sequence 182689, A
30	241	2.6	161	16 US-10-767-701-32672	Sequence 32672, A
31	202	2.2	585	15 US-10-424-599-285334	Sequence 285334,
32	188.5	2.0	2041	16 US-10-408-765A-1151	Sequence 1151, App
33	184	2.0	1979	14 US-10-205-823-419	Sequence 419, App
34	180.5	2.0	1413	15 US-10-369-493-6092	Sequence 6092, Ap
35	180	2.0	996	9 US-09-815-242-5251	Sequence 5251, Ap
36	180	2.0	1009	9 US-09-815-242-12141	Sequence 12141, A
37	179.5	1.9	2579	16 US-10-473-741-6	Sequence 6, Appli
38	179.5	1.9	2792	16 US-10-473-741-4	Sequence 4, Appli
39	179.5	1.9	2856	16 US-10-473-741-3	Sequence 3, Appli
40	178	1.9	1164	15 US-10-369-493-9770	Sequence 9770, Ap
41	177.5	1.9	2503	17 US-10-828-985A-11	Sequence 11, Appl
42	177.5	1.9	2543	17 US-10-828-985A-9	Sequence 9, Appli
43	177.5	1.9	2568	17 US-10-828-985A-7	Sequence 7, Appli
44	177.5	1.9	2816	15 US-10-240-145-145	Sequence 145, App
45	177	1.9	1151	16 US-10-408-765A-994	Sequence 994, App

ALIGNMENTS

RESULT 1

US-10-758-672A-2
; Sequence 2, Application US/10758672A
; Publication No. US20040185037A1
; GENERAL INFORMATION:
; APPLICANT: Han, et al.
; TITLE OF INVENTION: HUMAN E3 ALPHA UBIQUITIN LIGASE FAMILY
; FILE REFERENCE: 01017/35966B
; CURRENT APPLICATION NUMBER: US/10/758,672A
; PRIOR FILING DATE: 2004-01-15
; PRIOR APPLICATION NUMBER: US 09/724,126
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 60/187,911
; PRIOR FILING DATE: 2000-03-08
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 2
; LENGTH: 1749
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-758-672A-2

Query Match	100.0%;	Score	9224;	DB	16;	Length	1749;
Best Local Similarity	100.0%;	Pred. No.	0;				
Matches	1749;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	MADREAGGTERMEISAEIPQTPQRLASWDDQVDYTAFLHLHLAQLVPEIYFAEMDPDLE	60				
Db	1	MADREAGGTERMEISAEIPQTPQRLASWDDQVDYTAFLHLHLAQLVPEIYFAEMDPDLE	60				
Qy	61	KOESVQMSIPTLEWYLFGEEDPDICLEKLHSGAPQICGVRFKSGTTCRCAIDPT	120				
Db	61	KOESVQMSIPTLEWYLFGEEDPDICLEKLHSGAPQICGVRFKSGTTCRCAIDPT	120				

QY 121 CVLCMDCFODSVHKHRYRMHTSTGGPCDCGDTBAWKTGPFVNHGPRAGTIIKENSRC 180
DB 121 CVLCMDCFODSVHKHRYRMHTSTGGPCDCGDTBAWKTGPFVNHGPRAGTIIKENSRC 180
QY 181 PLNEEVIVQARKIPPSVIVYVEMTIIWEEKLPPELQIREKNERYCYLVFNDEHSHYDH 240
DB 181 PLNEEVIVQARKIPPSVIVYVEMTIIWEEKLPPELQIREKNERYCYLVFNDEHSHYDH 240
QY 241 VIYSQALDCELAELAHQHTTAIDKEGRRAVAGAYAAQCEAKEDIKSHSENVSHPLHV 300
DB 241 VIYSQALDCELAELAHQHTTAIDKEGRRAVAGAYAAQCEAKEDIKSHSENVSHPLHV 300
QY 301 EVLHSEIMAHQFALRLGSMWKNISYSDRFQIFCOACLRPEPSENCLISRLMLMDA 360
DB 301 EVLHSEIMAHQFALRLGSMWKNISYSDRFQIFCOACLRPEPSENCLISRLMLMDA 360
QY 361 KLYKGARKILHBLIFSSPFMEYKCLFAMEFVKYKQIQKEYISDDHDSISITALSQV 420
DB 361 KLYKGARKILHBLIFSSPFMEYKCLFAMEFVKYKQIQKEYISDDHDSISITALSQV 420
QY 421 MFTVPTLARHLIEBQNVISVITETLLEVLPEYLDNRNKNFQYSGDKLGRVYAVICDLK 480
DB 421 MFTVPTLARHLIEBQNVISVITETLLEVLPEYLDNRNKNFQYSGDKLGRVYAVICDLK 480
QY 481 YLILSKPTIWTBRLRMQFLEGRPSFLKILTCMOGMEIIRROYGOHIEVDPDWEAALAIQM 540
DB 481 YLILSKPTIWTBRLRMQFLEGRPSFLKILTCMOGMEIIRROYGOHIEVDPDWEAALAIQM 540
QY 541 QLKNIILMFQEWACADELLLVAYKECHKAVMRCSTSFISSTKVVSQSGHSLKTSYRV 600
DB 541 QLKNIILMFQEWACADELLLVAYKECHKAVMRCSTSFISSTKVVSQSGHSLKTSYRV 600
QY 601 SEDLSVILPLSRTLAGLHVRLSRGAVSRHLHFVSFEDFQVEVLVEYPLRCLVLVAQV 660
DB 601 SEDLSVILPLSRTLAGLHVRLSRGAVSRHLHFVSFEDFQVEVLVEYPLRCLVLVAQV 660
QY 661 AEWRRNGLSLSQVFPYQDVKCRBEMDKIIMLQIGASLMDPNKFLLLVLRVELABA 720
DB 661 AEWRRNGLSLSQVFPYQDVKCRBEMDKIIMLQIGASLMDPNKFLLLVLRVELABA 720
QY 721 FNKTISTKQDQLIKOYNTLIEEMQLVLIYIGERYVPGVGNVTKEBVTWREIHLICIEP 780
DB 721 FNKTISTKQDQLIKOYNTLIEEMQLVLIYIGERYVPGVGNVTKEBVTWREIHLICIEP 780
QY 781 MPHSAIAKLPNENNETGLENVINKVATPKPGVSGHGVYELKDBSLKDFNMYFHYSK 840
DB 781 MPHSAIAKLPNENNETGLENVINKVATPKPGVSGHGVYELKDBSLKDFNMYFHYSK 840
QY 841 TOHSAEHMOKKRRKQENKDEALPPPPPEPCFAPSKVINLNCDDIMMYILRTVPERAID 900
DB 841 TOHSAEHMOKKRRKQENKDEALPPPPPEPCFAPSKVINLNCDDIMMYILRTVPERAID 900
QY 901 TDSNLWTEGMLQWAFHILALGLLEKQQLQKAPBEEVTFDYHKAISRLGSSANNIOMLLE 960
DB 901 TDSNLWTEGMLQWAFHILALGLLEKQQLQKAPBEEVTFDYHKAISRLGSSANNIOMLLE 960
QY 961 KLKGPQLBGQKMDITWILQMFDTVRLREKSLIVATTSGSEIKNDBITDKKAEK 1020
DB 961 KLKGPQLBGQKMDITWILQMFDTVRLREKSLIVATTSGSEIKNDBITDKKAEK 1020
QY 1021 RXBAARLHROKIMQMSALQKNFIETHKLMYDNTSEMPGKEDSIWEEESTPAVSDYSRI 1080
DB 1021 RXBAARLHROKIMQMSALQKNFIETHKLMYDNTSEMPGKEDSIWEEESTPAVSDYSRI 1080
QY 1081 ALGPKRGPSVTEKEVLTCILQBEQEVKLENNAMVLSACVQKSTALTQHRGKPIELSGEA 1140
DB 1081 ALGPKRGPSVTEKEVLTCILQBEQEVKLENNAMVLSACVQKSTALTQHRGKPIELSGEA 1140
QY 1141 LDPLFMDPDLAYTGTYSCHGVHMAVCWQKYFEAVOLSSQQRHVDLFDLESGEVLCPLC 1200
DB 1141 LDPLFMDPDLAYTGTYSCHGVHMAVCWQKYFEAVOLSSQQRHVDLFDLESGEVLCPLC 1200
QY 1201 KSLCNTVPIPIPLQPKINSENADALQLLTLARWIQTVLARISGVNIRHAKGENPIPIF 1260

DB 1201 KSLCNTVPIPIPLQPKINSENADALQLLTLARWIQTVLARISGVNIRHAKGENPIPIF 1260
QY 1261 FNQMGDSTLBFHSLSPGVESSIKYNSIKEMVILFATTIYRIGLVKVPDERDRVPM 1320
DB 1261 FNQMGDSTLBFHSLSPGVESSIKYNSIKEMVILFATTIYRIGLVKVPDERDRVPM 1320
QY 1321 TWSCAFTIOAIENLGDGKPLFCALONROHNGIKALMQFAVAQRITCPOVLQKHLVR 1380
DB 1321 TWSCAFTIOAIENLGDGKPLFCALONROHNGIKALMQFAVAQRITCPOVLQKHLVR 1380
QY 1381 LLSVVLPNIKSEDTPCLLSIDLFHVLVGAFLAPSLYWDVDPVLOPSSVSSSYNHLYL 1440
DB 1381 LLSVVLPNIKSEDTPCLLSIDLFHVLVGAFLAPSLYWDVDPVLOPSSVSSSYNHLYL 1440
QY 1441 LITMAHMLQIILLTVDTGLPLAQVQEDSEAHASASFPABISQYTSGSIGCDIPGYLWVS 1500
DB 1441 LITMAHMLQIILLTVDTGLPLAQVQEDSEAHASASFPABISQYTSGSIGCDIPGYLWVS 1500
QY 1501 LKNGITPVLCALPFHYLLGVTPPELHTMSAEGYSALCSYLSLPTNLFLFOEYWD 1560
DB 1501 LKNGITPVLCALPFHYLLGVTPPELHTMSAEGYSALCSYLSLPTNLFLFOEYWD 1560
QY 1561 VRPLLQRCADPALINCLKQKNTVVRYPKKNLSIELPDDYSCLLNQASHFRCPSADDE 1620
DB 1561 VRPLLQRCADPALINCLKQKNTVVRYPKKNLSIELPDDYSCLLNQASHFRCPSADDE 1620
QY 1621 RKHPVLCLFCGAILCSQNICCOEIVNGBEVGACIFPHALHCGAGVCIPIKIRECRVVLVEG 1680
DB 1621 RKHPVLCLFCGAILCSQNICCOEIVNGBEVGACIFPHALHCGAGVCIPIKIRECRVVLVEG 1680
QY 1681 KARGCAYPAPYLDEYGETDGLKGNPLHLSRVRKHLVWQOHCIIIEETARSQETNOM 1740
DB 1681 KARGCAYPAPYLDEYGETDGLKGNPLHLSRVRKHLVWQOHCIIIEETARSQETNOM 1740
QY 1741 LFGFNWQLL 1749
DB 1741 LFGFNWQLL 1749

RESULT 2

US-10-357-819-10
; Sequence 10, Application US/10357819
; Publication No. US20040259774A1
; GENERAL INFORMATION:
; APPLICANT: Alvarez, Enrique
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Gerlach, Valerie L.
; APPLICANT: Gorman, Linda
; APPLICANT: Guo, Xiaojia
; APPLICANT: Ji, Weizhen
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Li, Li
; APPLICANT: Miller, Charles E.
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Patturajan, Meera
; APPLICANT: Rastelli, Luca
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-538A
; CURRENT APPLICATION NUMBER: US/10/357,819
; CURRENT FILING DATE: 2003-02-03
; PRIOR APPLICATION NUMBER: 09/520,781
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 09/584,411
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: 09/783,436
; PRIOR FILING DATE: 2001-02-14


```

; PRIOR APPLICATION NUMBER: 10/085,198
; PRIOR FILING DATE: 2002-02-25
; PRIOR APPLICATION NUMBER: 60/353,301
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/355,099
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/356,424
; PRIOR FILING DATE: 2002-02-12
; PRIOR APPLICATION NUMBER: 60/358,239
; PRIOR FILING DATE: 2002-02-20
; PRIOR APPLICATION NUMBER: 60/358,608
; PRIOR FILING DATE: 2002-02-21
; PRIOR APPLICATION NUMBER: 60/359,367
; PRIOR FILING DATE: 2002-02-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 142
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 10
; LENGTH: 1749
; TYPE: PR1
; ORGANISM: Homo sapiens
US-10-357-819-10

Query Match      100.0%; Score 9224; DB 16; Length 1749;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1749; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MADEERAGGTERMEISAELEPQTQRLASWDDQVDFYTAFLHHLAQIVPEIYFAEMDPDLE 60
DB 1 MADEERAGGTERMEISAELEPQTQRLASWDDQVDFYTAFLHHLAQIVPEIYFAEMDPDLE 60

QY 61 KQESVQMSIPTLEWYLFGEPPDICLEKLKHSAPQLCGRVFKSGETTYSCRDCAIDPT 120
DB 61 KQESVQMSIPTLEWYLFGEPPDICLEKLKHSAPQLCGRVFKSGETTYSCRDCAIDPT 120

QY 121 CVLQMDCFQDSVHKHRYKMHSTGGGFCDCGDTAEWKTPPCVNHPEGRAGTIKENSRC 180
DB 121 CVLQMDCFQDSVHKHRYKMHSTGGGFCDCGDTAEWKTPPCVNHPEGRAGTIKENSRC 180

QY 181 PLNEEVIVQARKIPPSVIKYVEMTWEEBKEPLPELQIREKNERYCYLVFNDHHSYDH 240
DB 181 PLNEEVIVQARKIPPSVIKYVEMTWEEBKEPLPELQIREKNERYCYLVFNDHHSYDH 240

QY 241 VIYSQALDCELAZEAQLHTTAIDKEGRAVAGAYAAQAEKEDIKSHSNVSHPLHV 300
DB 241 VIYSQALDCELAZEAQLHTTAIDKEGRAVAGAYAAQAEKEDIKSHSNVSHPLHV 300

QY 301 EVLHSEIMAHQKPAFLGSMNKNINSYSSDPQIFCOACLRPEPDSNPCLISRLMLWDA 360
DB 301 EVLHSEIMAHQKPAFLGSMNKNINSYSSDPQIFCOACLRPEPDSNPCLISRLMLWDA 360

QY 361 KLYKGARKILHELIFSSPFMEYKCLFAMEFVKYKQKQKEYISDDHRSISITALSVO 420
DB 361 KLYKGARKILHELIFSSPFMEYKCLFAMEFVKYKQKQKEYISDDHRSISITALSVO 420

QY 421 MFTVPTLARHLIEBQNVISVITETLEVLPEYLDNRNKNFQYSDQKLGRIYAVICDLK 480
DB 421 MFTVPTLARHLIEBQNVISVITETLEVLPEYLDNRNKNFQYSDQKLGRIYAVICDLK 480

QY 481 YILISKPTTWELRMQFLGFRSFLKILTCQGHBEIRROVQGHIEVDPDWEAAIAIQM 540
DB 481 YILISKPTTWELRMQFLGFRSFLKILTCQGHBEIRROVQGHIEVDPDWEAAIAIQM 540

QY 541 QLNKILLMFEQWACADELLLVAYKECHAVNRKSTSFSSKTVVQSGHSLKTSYRV 600
DB 541 QLNKILLMFEQWACADELLLVAYKECHAVNRKSTSFSSKTVVQSGHSLKTSYRV 600

QY 601 SEDLSVHLPLSRTLAGLHVLSRLGAVSRLEHFEVSFQVEVLVEYPLRCLVLVAQV 660
DB 601 SEDLSVHLPLSRTLAGLHVLSRLGAVSRLEHFEVSFQVEVLVEYPLRCLVLVAQV 660

QY 661 AEMWRNGLSLISQVFFYQDVVKREBMYDKDIIMLOIGASLMDPNKFLLLVLQRYELAEA 720

```

```

DB 661 AEMWRNGLSLISQVFFYQDVVKREBMYDKDIIMLOIGASLMDPNKFLLLVLQRYELAEA 720
QY 721 FNKTISTKQDQLIKQYNTLIEBMLQVLIYVGERVYVGVGNVTKBEVTMRREIHLCLCIEP 780
DB 721 FNKTISTKQDQLIKQYNTLIEBMLQVLIYVGERVYVGVGNVTKBEVTMRREIHLCLCIEP 780
QY 781 MPHSAIAKNLPENNENNETGLENVINKVATFKKPGVSGHGVYELKDBSLKDFNMYFYHYSK 840
DB 781 MPHSAIAKNLPENNENNETGLENVINKVATFKKPGVSGHGVYELKDBSLKDFNMYFYHYSK 840
QY 841 TOHSAKAEHQKRRKQENKDEALPPPPPEFPCAFSKVINLINCNDIMMYILTVPERAID 900
DB 841 TOHSAKAEHQKRRKQENKDEALPPPPPEFPCAFSKVINLINCNDIMMYILTVPERAID 900
QY 901 TDSNLWTEGMLQMAFHIALGALLEBEKQQLQKAPBEVTFDFYHKASRLGSSAMNQLMLE 960
DB 901 TDSNLWTEGMLQMAFHIALGALLEBEKQQLQKAPBEVTFDFYHKASRLGSSAMNQLMLE 960
QY 961 KLKGIPOLEGQKDMITWILQMPDVTVKRLREKSLIVATTSGSESTKNDBITHDKEKARK 1020
DB 961 KLKGIPOLEGQKDMITWILQMPDVTVKRLREKSLIVATTSGSESTKNDBITHDKEKARK 1020
QY 1021 RKAEARLHRQKIMQMSALQKNFIETHKLMYDNTSEMPGKEDSIMEBESTPAVDSYRI 1080
DB 1021 RKAEARLHRQKIMQMSALQKNFIETHKLMYDNTSEMPGKEDSIMEBESTPAVDSYRI 1080
QY 1081 ALGPKRGPSTVEKEVLTCTILCOBEQEVKIENNAMVLSACVQKSTALTQHRGKPIELSGEA 1140
DB 1081 ALGPKRGPSTVEKEVLTCTILCOBEQEVKIENNAMVLSACVQKSTALTQHRGKPIELSGEA 1140
QY 1141 LDPLFMDPDLAYGTYTSGCHVMHACVQKYPEAVOLSSQORIHVDLFDLESGEVLCPLC 1200
DB 1141 LDPLFMDPDLAYGTYTSGCHVMHACVQKYPEAVOLSSQORIHVDLFDLESGEVLCPLC 1200
QY 1201 KSLCNTVPIPIPLQPKINSENADALAOULLLARIQTVLARIQVNIHAKENPIPIF 1260
DB 1201 KSLCNTVPIPIPLQPKINSENADALAOULLLARIQTVLARIQVNIHAKENPIPIF 1260
QY 1261 FNQMGDSTLBFHSLTSGFVSESSIKYSIKEMVILPATIYIRIGLKVPPDPRVPM 1320
DB 1261 FNQMGDSTLBFHSLTSGFVSESSIKYSIKEMVILPATIYIRIGLKVPPDPRVPM 1320
QY 1321 TWSTCAFTIOAENILGDEGKPLFCALQNRQHNGIKALMQFAVQRICTPQVLIQKHLVR 1380
DB 1321 TWSTCAFTIOAENILGDEGKPLFCALQNRQHNGIKALMQFAVQRICTPQVLIQKHLVR 1380
QY 1381 LLSVVLVPLNKSEDTPLCLLSIDLPHVLVGAFLAPPSLYWDDPVDLQPSVSSSYNHLXLFH 1440
DB 1381 LLSVVLVPLNKSEDTPLCLLSIDLPHVLVGAFLAPPSLYWDDPVDLQPSVSSSYNHLXLFH 1440
QY 1441 LITMAHMLQILLITVDVTGLPLAQVQEDSEAHSAFFAEISQYTSGSIGCDIPGYLWVS 1500
DB 1441 LITMAHMLQILLITVDVTGLPLAQVQEDSEAHSAFFAEISQYTSGSIGCDIPGYLWVS 1500
QY 1501 LKNGITPYLRCAALPFHYLLGVTTPPEELHTNSAGEYSAALCSYLSLPTNLFLFQYWDY 1560
DB 1501 LKNGITPYLRCAALPFHYLLGVTTPPEELHTNSAGEYSAALCSYLSLPTNLFLFQYWDY 1560
QY 1561 VRPLLQRCADPALINCLKQNTVYVYPRKNSLIELEPDDYSCLLNQASHPRCPASADDE 1620
DB 1561 VRPLLQRCADPALINCLKQNTVYVYPRKNSLIELEPDDYSCLLNQASHPRCPASADDE 1620
QY 1621 RKHPVLCIFCGAILCSQNICCOEIVNGEVEGACIFHALHCGAGVCIPLKIRECRVVLVEG 1680
DB 1621 RKHPVLCIFCGAILCSQNICCOEIVNGEVEGACIFHALHCGAGVCIPLKIRECRVVLVEG 1680
QY 1681 KARGCAYPAPYLDVEGETDPLKRGKPNLHLSRERYKHLVWQHQHCTIEETARSQETQM 1740
DB 1681 KARGCAYPAPYLDVEGETDPLKRGKPNLHLSRERYKHLVWQHQHCTIEETARSQETQM 1740
QY 1741 LFGFNWQLL 1749
DB 1741 LFGFNWQLL 1749

```

RESULT 3									
US-10-758-636A-2									
; Sequence 2, Application US/10758636A									
; Publication No. US20050089876A1									
; GENERAL INFORMATION:									
; APPLICANT: Han, et al.									
; TITLE OF INVENTION: HUMAN E3 ALPHA UBIQUITIN LIGASE FAMILY									
; FILE REFERENCE: 01017/35966C									
; CURRENT APPLICATION NUMBER: US/10/758,636A									
; CURRENT FILING DATE: 2004-01-15									
; PRIOR APPLICATION NUMBER: US 09/724,126									
; PRIOR FILING DATE: 2000-11-28									
; PRIOR APPLICATION NUMBER: US 60/187,911									
; PRIOR FILING DATE: 2000-03-08									
; NUMBER OF SEQ ID NOS: 29									
; SOFTWARE: PatentIn version 3.2									
; SEQ ID NO 2									
; LENGTH: 1749									
; TYPE: PRT									
; ORGANISM: Homo sapiens									
US-10-758-636A-2									
Query Match 100.0%; Score 9224; DB 17; Length 1749;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 1749; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	MADEAGGTERMEISAELEPQTPQRLASWDDQVDFYAFLLHLAQLVPEIYPAE	MDPDL	60					
DB	1	MADEAGGTERMEISAELEPQTPQRLASWDDQVDFYAFLLHLAQLVPEIYPAE	MDPDL	60					
QY	61	KQESVQMSIFTPLEWYLGEPDPDLCLEKLSGAPQLCGRVFKSGETTYS	CRDCAIDPT	120					
DB	61	KQESVQMSIFTPLEWYLGEPDPDLCLEKLSGAPQLCGRVFKSGETTYS	CRDCAIDPT	120					
QY	121	CVLQMDCFQDSVHKHRYKWHSTGGGFCDCGDTAAWTKGPPCVNHEPGRAGT	IKENSR	180					
DB	121	CVLQMDCFQDSVHKHRYKWHSTGGGFCDCGDTAAWTKGPPCVNHEPGRAGT	IKENSR	180					
QY	181	PLNEEVIQARKIFPSVIKYVVEWMTIWEESKELEPPELQIREKNERYYCY	LFNDEHSHSYDH	240					
DB	181	PLNEEVIQARKIFPSVIKYVVEWMTIWEESKELEPPELQIREKNERYYCY	LFNDEHSHSYDH	240					
QY	241	VIYSLOALDCELAELAHQTHTAIDKEGRRAVAGAYAAQCBKAKEDIKSH	SENVSOHPLHV	300					
DB	241	VIYSLOALDCELAELAHQTHTAIDKEGRRAVAGAYAAQCBKAKEDIKSH	SENVSOHPLHV	300					
QY	301	EVLHSEIMAHQFALRLGSMWNKIMSYSSDFRQIFCOACLRPEPDS	ENPCLISRLMLWDA	360					
DB	301	EVLHSEIMAHQFALRLGSMWNKIMSYSSDFRQIFCOACLRPEPDS	ENPCLISRLMLWDA	360					
QY	361	KLYKGARKILHELIFSSFFMEWEYKGLFAMEFVKYKQLQKEYISDDH	DRSISITALS	420					
DB	361	KLYKGARKILHELIFSSFFMEWEYKGLFAMEFVKYKQLQKEYISDDH	DRSISITALS	420					
QY	421	MFTVPTLARHLIEBQNVISVITETLLLEVLPEYLDNRNKNFQGYSD	QKLGRVYAVICDLK	480					
DB	421	MFTVPTLARHLIEBQNVISVITETLLLEVLPEYLDNRNKNFQGYSD	QKLGRVYAVICDLK	480					
QY	481	YILISKPTITWTELRMQPLGFRSFLKILTCMQGMEIRROVGOHIE	VPDNEAAIAIQM	540					
DB	481	YILISKPTITWTELRMQPLGFRSFLKILTCMQGMEIRROVGOHIE	VPDNEAAIAIQM	540					
QY	541	QLKNILLMPQEWACDEBELLVAYKECHKAVMRCSTSFISSSKTVV	QSGHSLSTKSYRV	600					
DB	541	QLKNILLMPQEWACDEBELLVAYKECHKAVMRCSTSFISSSKTVV	QSGHSLSTKSYRV	600					
QY	601	SEDLVSIHPLSLTLAGLVRLSRLGAVSRLEHFEVDFQVEVLVEY	PLRCLVLVAQVV	660					
DB	601	SEDLVSIHPLSLTLAGLVRLSRLGAVSRLEHFEVDFQVEVLVEY	PLRCLVLVAQVV	660					
QY	661	AEWRRNGLSLISQVFFYQDVKCREEMYDKDIIMLQIGASLMDPNK	FLLLVLQRYELAEA	720					

Db 1741 LFGFNWQLL 1749

RESULT 4

US-10-287-218-1

Sequence 1, Application US/10287218

Publication No. US20030198975A1

GENERAL INFORMATION:

APPLICANT: INCYTE GENOMICS, INC.

APPLICANT: AZIMZAI, Yalda; AU-YOUNG, Janice K.

APPLICANT: BATRA, Sajeev; BAUGHN, Mariah R.

APPLICANT: BECHA, Shanya D.; BOROWSKY, Mark L.

APPLICANT: BUFORD, Neil; DING, Li

APPLICANT: ELLIOTT, Vicki S.; EMERLING, Brooke M.

APPLICANT: GANDHI, Ameena R.; GIBTZEN, Kimberly J.

APPLICANT: GRIFFIN, Jennifer A.; HAPALIA, April J.A.

APPLICANT: HONCHELL, Cynthia D.; LAL, Preeti G.

APPLICANT: LEE, Soo Yeun; LU, Dyung Aina M.

APPLICANT: ARVIZU, Chandra S.; RAMKUMAR, Jayalaxmi

APPLICANT: REDDY, Roopa; SANJANWALA, Madhu, M.

APPLICANT: TANG, Y. Tom; WALIA, Narinder K.

APPLICANT: WANG, Yu-mei, E.; WARREN, Bridget A.

APPLICANT: XU, Yuming; YANG, Junning

APPLICANT: YAO, Monique G.; YUE, Henry

APPLICANT: ZEBARJANIAN, Yeganeh

TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH

FILE REFERENCE: PI-0417 USA

CURRENT APPLICATION NUMBER: US/10/287,218

PRIOR FILING DATE: 2002-10-31

PRIOR APPLICATION NUMBER: PCT/US02/11152

PRIOR FILING DATE: 2002-04-05

PRIOR APPLICATION NUMBER: US 60/349,705

PRIOR FILING DATE: 2002-01-15

PRIOR APPLICATION NUMBER: US 60/295,263

PRIOR FILING DATE: 2001-06-01

PRIOR APPLICATION NUMBER: US 60/295,340

PRIOR FILING DATE: 2001-06-01

PRIOR APPLICATION NUMBER: US 60/293,727

PRIOR FILING DATE: 2001-05-25

PRIOR APPLICATION NUMBER: US 60/291,846

PRIOR FILING DATE: 2001-05-18

PRIOR APPLICATION NUMBER: US 60/291,662

PRIOR FILING DATE: 2001-05-16

PRIOR APPLICATION NUMBER: US 60/287,228

PRIOR FILING DATE: 2001-04-27

PRIOR APPLICATION NUMBER: US 60/286,820

PRIOR FILING DATE: 2001-04-26

PRIOR APPLICATION NUMBER: US 60/283,294

PRIOR FILING DATE: 2001-04-11

Remaining Prior Application data removed - See File Wrapper or PALM.

NUMBER OF SEQ ID NOS: 42

SOFTWARE: PERL Program

SEQ ID NO 1

LENGTH: 1738

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

OTHER INFORMATION: Incyte ID No. US20030198975A1 1351608CD1

US-10-287-218-1

Query Match 99.4%; Score 9168; DB 14; Length 1738;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 1738; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 MEISAEIPOTPORLASWDDQVDFTYAFHLHQAQLVPEIYPAEMDPDLKQESVQMSIP 71

Db 1 MEISAEIPOTPORLASWDDQVDFTYAFHLHQAQLVPEIYPAEMDPDLKQESVQMSIP 60

QY 72 TPLEWLFGEEDPDICLEKLGKSGAFQLCGRVFKSGRTTYSRDCDAIDPTCVLCMDCFQDS 131

Db 61 TPLEWLFGEEDPDICLEKLGKSGAFQLCGRVFKSGRTTYSRDCDAIDPTCVLCMDCFQDS 120

QY 132 VHKHRYXOMHTSTGCGFCDCGTBAWKTGPPCVNHEPGRAGTIKENSRCPLNEEVIQAR 191

Db 121 VHKHRYXOMHTSTGCGFCDCGTBAWKTGPPCVNHEPGRAGTIKENSRCPLNEEVIQAR 180

QY 192 KIFPSVIVYVEMTIIWEEKELPPELOIREKNERYCVLFNDEHSHSDHVIYSLQALDC 251

Db 181 KIFPSVIVYVEMTIIWEEKELPPELOIREKNERYCVLFNDEHSHSDHVIYSLQALDC 240

QY 252 ELAEQAQLHTTAIDKEGRRAVAGAAQCEAKEDIKSHSENVSOPLHVEVLHSEIHAHQ 311

Db 241 ELAEQAQLHTTAIDKEGRRAVAGAAQCEAKEDIKSHSENVSOPLHVEVLHSEIHAHQ 300

QY 312 KPALRLGSMWNKIMSYSSDFROIPOACLRBPDPSENPCILSRMLMDAKLKGARKILH 371

Db 301 KPALRLGSMWNKIMSYSSDFROIPOACLRBPDPSENPCILSRMLMDAKLKGARKILH 360

QY 372 ELIFSSPFMEYKCLFAMEFVKYKQKQKEYISDDHRSISITALSVMQPTVPTLARHL 431

Db 361 ELIFSSPFMEYKCLFAMEFVKYKQKQKEYISDDHRSISITALSVMQPTVPTLARHL 420

QY 432 IBEQNVISVITETLLLEVLPEYLDNRNKNFQGSQDKLGRVYAVICDLKYILISKPTWT 491

Db 421 IBEQNVISVITETLLLEVLPEYLDNRNKNFQGSQDKLGRVYAVICDLKYILISKPTWT 480

QY 492 ERLRQPLEGFRSFLKILTCQCMEEIRROVQGHIEVDPDWEAAIAIQMLKNILLMFOE 551

Db 481 ERLRQPLEGFRSFLKILTCQCMEEIRROVQGHIEVDPDWEAAIAIQMLKNILLMFOE 540

QY 552 WCACDEBELLVAYKECHKAVMRCSTPISSSKTVVQSCGHSLETYSRVSEDLVSHLPL 611

Db 541 WCACDEBELLVAYKECHKAVMRCSTPISSSKTVVQSCGHSLETYSRVSEDLVSHLPL 600

QY 612 SRTLAGLHVRLSRLGAVSRLEHFVSFEDFQVEVLVEYPLRCLVLVAQVVAEMRRENGLSL 671

Db 601 SRTLAGLHVRLSRLGAVSRLEHFVSFEDFQVEVLVEYPLRCLVLVAQVVAEMRRENGLSL 660

QY 672 ISQVFFYQDVKCREEMDYDKIIMLQIGASLMDPNKFLLLVLQRYELAEAFNKTISTKQD 731

Db 661 ISQVFFYQDVKCREEMDYDKIIMLQIGASLMDPNKFLLLVLQRYELAEAFNKTISTKQD 720

QY 732 LIKQNTLIEBMLQVLIYVGERVYVGVNVTKEBVTWREIHLACIEPMHSAIAKNLP 791

Db 721 LIKQNTLIEBMLQVLIYVGERVYVGVNVTKEBVTWREIHLACIEPMHSAIAKNLP 780

QY 792 ENENNETGLENVINKVATFKPGVSGHGYELKQESLKDFFNYFYHYSKTOHSHKAEHMQK 851

Db 781 ENENNETGLENVINKVATFKPGVSGHGYELKQESLKDFFNYFYHYSKTOHSHKAEHMQK 840

QY 852 KRRKOENKDEALPPPPPPPCFAPSKVINLNCIDIMMVLRTVFFERADTDSNLTWTEGML 911

Db 841 KRRKOENKDEALPPPPPPPCFAPSKVINLNCIDIMMVLRTVFFERADTDSNLTWTEGML 900

QY 912 QMAFHILALGLLEKQLOKAPREBVTDFVHKASRLGSSAWNOMLEKLGIPQESQ 971

Db 901 QMAFHILALGLLEKQLOKAPREBVTDFVHKASRLGSSAWNOMLEKLGIPQESQ 960

QY 972 KDMITWILQMFDTVKRLREKSLIVATTSGSESKNDEITHDKKAEKRAEAAARLHRQ 1031

Db 961 KDMITWILQMFDTVKRLREKSLIVATTSGSESKNDEITHDKKAEKRAEAAARLHRQ 1020

QY 1032 KIMAQMSALQKNFIETHKLMYDNTSEMPGKSDSIMEESTPAVSYSRIALGPGRGPSVT 1091

Db 1021 KIMAQMSALQKNFIETHKLMYDNTSEMPGKSDSIMEESTPAVSYSRIALGPGRGPSVT 1080

QY 1092 EKEVLCTILCOBQEVKLENNAMVLSACVQKSTALTOHRGKPIELSGALDPLFMDPDLA 1151

Db 1081 EKEVLCTILCOBQEVKLENNAMVLSACVQKSTALTOHRGKPIELSGALDPLFMDPDLA 1140

QY 1152 YGTYTSGCHVHAWCWOKYPEAVQLSSQQRHIVDLFDLESGEVLCPLCKSLCNTVPIII 1211

Db 1141 YGTYTSGCHVHAWCWOKYPEAVQLSSQQRHIVDLFDLESGEVLCPLCKSLCNTVPIII 1200

QY 1212 PLQPOKINSNADALAQLLTLARWITQTVLARISGVNIRHAKGENPIPIFFNQMGDSTLE 1271

Db 1201 PLOPQKINSENADALQALLTLARWLTQVTLARISGNIRHAKGENPIPIFFNQGMDSTLE 1260
QY 1272 FHSILSPGVSESSIKYSNISKEMVILFATTIYIRIGLKVPDPDRPRVPMLTWSTCAFTTQA 1331
Db 1261 FHSILSPGVSESSIKYSNISKEMVILFATTIYIRIGLKVPDPDRPRVPMLTWSTCAFTTQA 1320
QY 1332 IENLLGDEKPLFGALONRQHGLKALMQPAAQRTCPQVLIQKHLVRLSVLPNTKS 1391
Db 1321 IENLLGDEKPLFGALONRQHGLKALMQPAAQRTCPQVLIQKHLVRLSVLPNTKS 1380
QY 1392 EDTPCLLSIDLFAVLGAVLAPPSLYWDDPVDLPQSSVSSSYNHLYLPHLITMAHMLQIL 1451
Db 1381 EDTPCLLSIDLFAVLGAVLAPPSLYWDDPVDLPQSSVSSSYNHLYLPHLITMAHMLQIL 1440
QY 1452 LTVDTGLPLAQVOEDSEEAHSAASSPFAETISQVTSIGSDIPGWYLMVSLKNGITPYLRC 1511
Db 1441 LTVDTGLPLAQVOEDSEEAHSAASSPFAETISQVTSIGSDIPGWYLMVSLKNGITPYLRC 1500
QY 1512 AALPHYLLGVTPPELHTNSAGEYSALCSYLSLPTNLLPLPOBYWDTVRLLQRCAD 1571
Db 1501 AALPHYLLGVTPPELHTNSAGEYSALCSYLSLPTNLLPLPOBYWDTVRLLQRCAD 1560
QY 1572 PALLNCLKQKNTVRYPRKRNLSIELPDDYSCLLQASHFRCPRSADDERKHPVLCFCG 1631
Db 1561 PALLNCLKQKNTVRYPRKRNLSIELPDDYSCLLQASHFRCPRSADDERKHPVLCFCG 1620
QY 1632 AILCSQNICCOBIVNGEEVGACIFHALHCGAGVCIFLKIRECRVVLVEGKARGCAYPAPY 1691
Db 1621 AILCSQNICCOBIVNGEEVGACIFHALHCGAGVCIFLKIRECRVVLVEGKARGCAYPAPY 1680
QY 1692 LDEYGETDPLKRGKGNPLHLRSRYRKLHLVWQOHCIIIEIARSQETNOMLFGFNWOLL 1749
Db 1681 LDEYGETDPLKRGKGNPLHLRSRYRKLHLVWQOHCIIIEIARSQETNOMLFGFNWOLL 1738

RESULT 5

US-10-474-291-1
; Sequence 1, Application US/10474291
; Publication No. US20040132043A1
; GENERAL INFORMATION:
; APPLICANT: AZIMZAI, Yalda; AU-YOUNG, Janice K.
; APPLICANT: BATRA, Sajeev; BAUGHN, Mariah R.
; APPLICANT: BECHA, Shanya D.; BOWSKY, Mark L.
; APPLICANT: BURFORD, Neil; DING, Li
; APPLICANT: ELLIOTT, Vicki S.; EMERLING, Brooke M.
; APPLICANT: GANDHI, Ameena R.; GIETZEN, Kimberly J.
; APPLICANT: GRIFFIN, Jennifer A.; HAPALIA, April J.A.
; APPLICANT: HONCHELL, Cynthia D.; LAL, Preeti G.
; APPLICANT: LEE, Soo Yeun; LU, Dyung Aina M.
; APPLICANT: ARVIZU, Chandra S.; RAMKUMAR, Jayalaxmi
; APPLICANT: REDDY, Roopa M.; SANJANWALA, Madhusudan M.
; APPLICANT: TANG, Y. Tom; CHAWLA, Navinder K.
; APPLICANT: WANG, Yu-Mei E.; WARREN, Bridget A.
; APPLICANT: XU, Yuming; YANG, Junming
; APPLICANT: YAO, Monique G.; YUE, Henry
; APPLICANT: ZEBARJADIAN, Yeganeh
; TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH
; FILE REFERENCE: PI-0417 USN
; CURRENT APPLICATION NUMBER: US/10/474,291
; CURRENT FILING DATE: 2003-10-06
; PRIOR APPLICATION NUMBER: PCT/US02/11152
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: US 60/349,705
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: US 60/295,263
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/295,340
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/293,727
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/291,846
; PRIOR FILING DATE: 2001-05-18

; PRIOR APPLICATION NUMBER: US 60/291,662
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: US 60/287,228
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/286,820
; PRIOR FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: US 60/283,294
; PRIOR FILING DATE: 2001-04-11
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 1738
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 1351608CD1
US-10-474-291-1

Query Match 99.4%; Score 9168; DB 16; Length 1738;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1738; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 12 MEISAEPLPOTQRLASWMDQVDFYTAFLHHLAQLVPEIYFAEMDPDLKOEESVQMSIF 71
Db 1 MEISAEPLPOTQRLASWMDQVDFYTAFLHHLAQLVPEIYFAEMDPDLKOEESVQMSIF 60
QY 72 TPLEWYLFGEPPDICLEKLKHSQAPQLCGRVFKSGETTYSCRDCAIDPTCVLCMDCFQDS 131
Db 61 TPLEWYLFGEPPDICLEKLKHSQAPQLCGRVFKSGETTYSCRDCAIDPTCVLCMDCFQDS 120
QY 132 VHKHRYKMHSTTGGFCDCGDTAUKTGPVNHPEGRAGTIKENSRCPLNEEVIQAR 191
Db 121 VHKHRYKMHSTTGGFCDCGDTAUKTGPVNHPEGRAGTIKENSRCPLNEEVIQAR 180
QY 192 KIPPSVIKYVVMETIWESEKELPPELOIREKNERYCYLVFNDEHHSYDHVYISLQALDC 251
Db 181 KIPPSVIKYVVMETIWESEKELPPELOIREKNERYCYLVFNDEHHSYDHVYISLQALDC 240
QY 252 ELAERLAQLHTTAIDKEGRRAVAGAAQCEAKEDIKSHSENVSHPLHVEVLHSGEIMAHQ 311
Db 241 ELAERLAQLHTTAIDKEGRRAVAGAAQCEAKEDIKSHSENVSHPLHVEVLHSGEIMAHQ 300
QY 312 KPALRLGSMWNKIMSYSSDFRQIFCOACLRREPDPSENCLISRLMLWDAKLYKGARKILH 371
Db 301 KPALRLGSMWNKIMSYSSDFRQIFCOACLRREPDPSENCLISRLMLWDAKLYKGARKILH 360
QY 372 ELIFSSPFMEYKCLFAMEFVKYKQLOKEYISDDHRSISITALSVMQFTVPTLARHL 431
Db 361 ELIFSSPFMEYKCLFAMEFVKYKQLOKEYISDDHRSISITALSVMQFTVPTLARHL 420
QY 432 IEQNVISVITETLLEVLPEYLDNRNKNFNGYQSKLGRVYAVICDLKYILISKPTIWT 491
Db 421 IEQNVISVITETLLEVLPEYLDNRNKNFNGYQSKLGRVYAVICDLKYILISKPTIWT 480
QY 492 ERLRMQFLGFRSFLKILTCMOGHEEIPROVQOHLVDPDWEAAIAIQMLKNILLMFOE 551
Db 481 ERLRMQFLGFRSFLKILTCMOGHEEIPROVQOHLVDPDWEAAIAIQMLKNILLMFOE 540
QY 552 WCACDEELLVAYKECHKAVMRCSTSFISSTKTVVQSCGHSLETYSKYSVSDLSIHLPL 611
Db 541 WCACDEELLVAYKECHKAVMRCSTSFISSTKTVVQSCGHSLETYSKYSVSDLSIHLPL 600
QY 612 SRTLGLHVLRLSRLGAVSRLEHFVSFEDQVEVLVEYPLRCLVLVAQVVAEMWRNGLSL 671
Db 601 SRTLGLHVLRLSRLGAVSRLEHFVSFEDQVEVLVEYPLRCLVLVAQVVAEMWRNGLSL 660
QY 672 ISQVPIYQDVKCRBEMVDKDIIMLQIGASLMDPNKFLLLVLQRYELAEAFNKTIKTDOD 731
Db 661 ISQVPIYQDVKCRBEMVDKDIIMLQIGASLMDPNKFLLLVLQRYELAEAFNKTIKTDOD 720
QY 732 LIKQVNTLIBEMQLVLIYIGERYVPGVGNVTKBEVTWREIHLICIBFMPHSAIAKNLP 791

Db 721 LIKQYNTLIEMLQVLIYVGERYVGVGNVTKBEVTMRBIHLHLCIEBPHSALIAKNLP 780
QY 792 ENENNETGLENVINKVATFKPGVSGHGVYELKDESXKDPFNMVYFHYSKTQHSKAEHMQK 851
Db 781 ENENNETGLENVINKVATFKPGVSGHGVYELKDESXKDPFNMVYFHYSKTQHSKAEHMQK 840
QY 852 KRRQENKDEALPPPPPPPCFAPSKVINLLNCIDIMMYILRTVFERAIDTDSNLTGML 911
Db 841 KRRQENKDEALPPPPPPPCFAPSKVINLLNCIDIMMYILRTVFERAIDTDSNLTGML 900
QY 912 QNAFHILALGLLEKQLOKAPBEVTFDYFKASRLGSSANNIOMLLEKLGIPQLEQ 971
Db 901 QNAFHILALGLLEKQLOKAPBEVTFDYFKASRLGSSANNIOMLLEKLGIPQLEQ 960
QY 972 KOMITWILQWFTVTKRLREKSLIVATTSGESIKNDEITHDKAEARKKAEARLHRQ 1031
Db 961 KOMITWILQWFTVTKRLREKSLIVATTSGESIKNDEITHDKAEARKKAEARLHRQ 1020
QY 1032 KIMAQMSALQKQNFIEHKLKYNDTSEMPGKEDSIMEESTPAVDYSRIALGPGRGPSVT 1091
Db 1021 KIMAQMSALQKQNFIEHKLKYNDTSEMPGKEDSIMEESTPAVDYSRIALGPGRGPSVT 1080
QY 1092 EKEVLTCILCOBEOVKIENNAWLSACVOKSTALTQHRGKPTIELSGEALDPLFMDPDLA 1151
Db 1081 EKEVLTCILCOBEOVKIENNAWLSACVOKSTALTQHRGKPTIELSGEALDPLFMDPDLA 1140
QY 1152 YGTYTSGCGHVMHACVQKYFEAVQLSSQORIHVDLPDLESEGYLCPCKSLCNTVPII 1211
Db 1141 YGTYTSGCGHVMHACVQKYFEAVQLSSQORIHVDLPDLESEGYLCPCKSLCNTVPII 1200
QY 1212 PIQOKINSENADALAQLLTLARWIQTVLARIISGVNIRHAKENPIPIPFNGMGDSTLE 1271
Db 1201 PIQOKINSENADALAQLLTLARWIQTVLARIISGVNIRHAKENPIPIPFNGMGDSTLE 1260
QY 1272 FHSILSPGVESISKYSNISKEMVILPATIYRIGLKVPPDRDPVPMLTWTSCAFTIOA 1331
Db 1261 FHSILSPGVESISKYSNISKEMVILPATIYRIGLKVPPDRDPVPMLTWTSCAFTIOA 1320
QY 1332 IENLLGDEGKPLFGALQNRQHGLKALMQFAVAQRITCPQVLTQKHLVRLLSVLPNKS 1391
Db 1321 IENLLGDEGKPLFGALQNRQHGLKALMQFAVAQRITCPQVLTQKHLVRLLSVLPNKS 1380
QY 1392 EDTPCLLSIDLPHVLVGAVALPSPSYWDDPDLQPSVSSVSSNNHLYLPHLITMAHMLQIL 1451
Db 1381 EDTPCLLSIDLPHVLVGAVALPSPSYWDDPDLQPSVSSVSSNNHLYLPHLITMAHMLQIL 1440
QY 1452 LTVDTCPLAQOEDSEEAHSAASPFAETISQYTSIGCDIPGWYLVSLKNGITPYLRC 1511
Db 1441 LTVDTCPLAQOEDSEEAHSAASPFAETISQYTSIGCDIPGWYLVSLKNGITPYLRC 1500
QY 1512 AALFPHYLGVTPPEELHTNSABGEYSALCSYLSLPTNLPLFOEYWDVTRPLLQWCAD 1571
Db 1501 AALFPHYLGVTPPEELHTNSABGEYSALCSYLSLPTNLPLFOEYWDVTRPLLQWCAD 1560
QY 1572 PALLNCLKQKNTVVRYPKRNLSIELPDDYSCLLQAASHFRCPRSADDERKHPVLCFG 1631
Db 1561 PALLNCLKQKNTVVRYPKRNLSIELPDDYSCLLQAASHFRCPRSADDERKHPVLCFG 1620
QY 1632 AILCSQNTCCQIBVNGEEVGACIFALHCGAGVCIFLKIREDRVVLVEGKARGCAYPAPY 1691
Db 1621 AILCSQNTCCQIBVNGEEVGACIFALHCGAGVCIFLKIREDRVVLVEGKARGCAYPAPY 1680
QY 1692 LDEYGETDPLKRGKGNPLHLSRERYRKLHLVWQOHCIIIEIARSQETNOMLFGFNWQLL 1749
Db 1681 LDEYGETDPLKRGKGNPLHLSRERYRKLHLVWQOHCIIIEIARSQETNOMLFGFNWQLL 1738

RESULT 6

US-10-758-672A-19

; Sequence 19, Application US/10758672A

; Publication No. US20040185037A1

; GENERAL INFORMATION:

; APPLICANT: Han, et al.
; TITLE OF INVENTION: HUMAN E3 ALPHA UBIQUITIN LIGASE FAMILY
; FILE REFERENCE: 01017/35966B
; CURRENT APPLICATION NUMBER: US/10/758,672A
; CURRENT FILING DATE: 2004-01-15
; PRIOR APPLICATION NUMBER: US 09/724,126
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 60/187,911
; PRIOR FILING DATE: 2000-03-08
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 19
; LENGTH: 1734
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-758-672A-19

Query Match 98.5%; Score 9087; DB 16; Length 1734;

Best Local Similarity 98.8%; Pred. No. 0;

Matches 1728; Conservative 0; Mismatches 1; Indels 20; Gaps 1;

QY 1 MADEEAGGTGTERMEISAEIPQTPQRLASWDDQVDFYTAFLHHLAQLVPEIYFAEMDPDLE 60
Db 6 MADEEAGGTGTERMEISAEIPQTPQRLASWDDQVDFYTAFLHHLAQLVPEIYFAEMDPDLE 65
QY 61 KOEEVQMSIFTPLEWYLFGBDDPDIKLEKLKHSAGAPQLCGRVFKSGETTYSRCDCAIDPT 120
Db 66 KOEEVQMSIFTPLEWYLFGBDDPDIKLEKLKHSAGAPQLCGRVFKSGETTYSRCDCAIDPT 125
QY 121 CVLCHMDCFQDSVHKHRYKMHSTGCGFCDCGDTBAWTKGPPCVNHHFGRAGTIIKENSRC 180
Db 126 CVLCHMDCFQDSVHKHRYKMHSTGCGFCDCGDTBAWTKGPPCVNHHFGRAGTIIKENSRC 185
QY 181 PLNEEVIQVQARKIPFSVIKYVEMTMEEEKELPELQIREKNERYCYVLFNDEHHSYDH 240
Db 186 PLNEEVIQVQARKIPFSVIKYVEMTMEEEKELPELQIREKNERYCYVLFNDEHHSYDH 245
QY 241 VIYSLQALDCELAQAQLHTTAIDKEGRRVAKAGAYAAQCEAKEDIKHSSENVSOHPLHV 300
Db 246 VIYSLQALDCELAQAQLHTTAIDKEGRRVAKAGAYAAQCEAKEDIKHSSENVSOHPLHV 305
QY 301 EYLVSEIHAHQFALRGSSWNKINSYSSDPRQIFCOACLRPEPSENPCILSRMLWDA 360
Db 306 EYLVSEIHAHQFALRGSSWNKINSYSSDPRQIFCOACLRPEPSENPCILSRMLWDA 365
QY 361 KLYGARKILHELIPSSPFMEYKCLPAMEFVKYKQLEKEYISDDHDSISITALSVO 420
Db 366 KLYGARKILHELIPSSPFMEYKCLPAMEFVKYKQLEKEYISDDHDSISITALSVO 425
QY 421 MFTVPTLARHLIEBQNVISVITETLLEVLPEYLDNRNKNFQGYSDQKLGRIYAVICDLK 480
Db 426 MFTVPTLARHLIEBQNVISVITETLLEVLPEYLDNRNKNFQGYSDQKLGRIYAVICDLK 485
QY 481 YILISKPTIWTRELRMQFLEGFRSFLKLTTCMOGHEEIRROVQGHIEYVDPWEAAIAIQM 540
Db 486 YILISKPTIWTRELRMQFLEGFRSFLKLTTCMOGHEEIRROVQGHIEYVDPWEAAIAIQM 545
QY 541 QLNKILLMFQSCACDEBELLVAYKECHKAVMRCTSFISSTKVQSCGHSLETYSRV 600
Db 546 QLNKILLMFQSCACDEBELLVAYKECHKAVMRCTSFISSTKVQSCGHSLETYSRV 605
QY 601 SEDLVS IHLPLSRTLGLHVLRLSRLGAVSRHLHEFVSFEDFQVLEVEYPLRCLVLVAQV 660
Db 606 SEDLVS IHLPLSRTLGLHVLRLSRLGAVSRHLHEFVSFEDFQVLEVEYPLRCLVLVAQV 665
QY 661 AEMWRNGLSLISQVFFYQDVKCREMYDKDIIMLQIGASLMDPNKFLLLVLQRYELABA 720
Db 666 AEMWRNGLSLISQVFFYQDVKCREMYDKDIIMLQIGASLMDPNKFLLLVLQRYELABA 725
QY 721 FNKTISTKQDQLIKQYNTLIEMLQVLIYVGERYVGVGNVTKBEVTMRBIHLHLCIEP 780
Db 726 FNKTISTKQDQLIKQYNTLIEMLQVLIYVGERYVGVGNVTKBEVTMRBIHLHLCIEP 785

781 MPHSAIAKXLPENNETGLENVINKVATFKKPGVSGHGVYELKDESKDFMFPYHYSK 840
 786 MPHSAIAKXLPENNETGLENVINKVATFKKPGVSGHGVYELKDESKDFMFPYHYSK 845
 841 TQHSKAEHQKRRKQENKDEALPPPPPPFAPSKVINLNCDDIMYILRTVPERAID 900
 846 TQHSKAEHQKRRKQENKDEALPPPPPPFAPSKVINLNCDDIMYILRTVPERAID 905
 901 TDSNLWTEGMLQMAFHIALGELLEKQOLQKAPPEEVTFDFVHKASRLGSSAMNOMLLE 960
 906 TDSNLWTEGMLQMAFHIALGELLEKQOLQKAPPEEVTFDFVHKASRLGSSAMNOMLLE 965
 961 KLGKIPQLEGQKDMITWILQMDTVKRLREKSLCVATTSSESINKDEITHDKKARK 1020
 966 KLGKIPQLEGQKDMITWILQMDTVKRLREKSLCVATTSSESINKDEITHDKKARK 1025
 1021 RKAEARLHRQKIMQMSALQKNFIETHKLMYDNTSEMPGKEDSIMEESTPAVSDYSRI 1080
 1026 RKAEARLHRQKIMQMSALQKNFIETHKLMYDNTSEMPGKEDSIMEESTPAVSDYSRI 1085
 1081 ALGPKRGPSVTEKEVLTCLCOEBORVKIENNAVLACVQKSTALTQHRGKPIELSGEA 1140
 1086 ALGPKRGPSVTEKEVLTCLCOEBORVKIENNAVLACVQKSTALTQHRGKPIELSGEA 1145
 1141 LDPLFMDPLAYGTYTSGCGHVMHACWQKYFEAVQLSSQQRHVDLPDLESGEYLCLPC 1200
 1146 LDPLFMDPLAYGTYTSGCGHVMHACWQKYFEAVQLSSQQRHVDLPDLESGEYLCLPC 1205
 1201 KSLCNTVPIIPILOPKINSENADALQALLTLARMIQTVLARISGYNIRHAKGENPIPIF 1260
 1206 KSLCNTVPIIPILOPKINSENADALQALLTLARMIQTVLARISGYNIRHAKGENPIPIF 1265
 1261 FNQGMGDSLEPHSILSPGVESIKYNSIKEMVILPATTIYRIGLKVPDPDRDPRVPM 1320
 1266 FNQGMGDSLEPHSILSPGVESIKYNSIKEMVILPATTIYRIGLKVPDPDRDPRVPM 1325
 1321 TWSTCAFTTQALLENLGDGKPLFGALQNRQNGKALMQPAAQRITCPQVLIQKHLVR 1380
 1326 TWSTCAFTTQALLENLGDGKPLFGALQNRQNGKALMQPAAQRITCPQVLIQKHLVR 1385
 1381 LLSVLPNPKSDETPCLLSIDLPHVLVGAVALAPPSLYMDPVDLPQSSVSSSYNHLYLPH 1440
 1386 LLSVLPNPKSDETPCLLSIDLPHVLVGAVALAPPSLYMDPVDLPQSSVSSSYNHLYLPH 1445
 1441 LITMAHMLQILTVDTGLPLAQVEDSEAHGASFPABISQYTSIGCDIPGHWLWS 1500
 1446 LITMAHMLQILTVDTGLPLAQVEDSEAHGASFPABISQYTSIGCDIPGHWLWS 1505
 1501 LKNGITPYLRCAALPHYLLGVTPPELHTNABGEYSALCSYLSLPTNLFLFLFOBYMDT 1560
 1506 LKNGITPYLRCAALPHYLLGVTPPELHTNABGEYSALCSYLSLPTNLFLFLFOBYMDT 1565
 1561 VRPLLQRACADPALLNCLKQKNTVVYPRKRNLSIELPDDYSCLLNQASHFRCPRSADDE 1620
 1566 VRPLLQRACADPALLNCLKQKNTVVYPRKRNLSIELPDDYSCLLNQASHFRCPRSADDE 1625
 1621 RHPVLCLFCGATLCSQNTCCQBIIVNGRBYGACIFHAIHCGAGVCIFLIRECRVVLVEG 1680
 1626 RHPVLCLFCGATLCSQNTCCQBIIVNGRBYGACIFHAIHCGAGVCIFLIRECRVVLVEG 1685
 1681 KARGCAYPAPYLDYGETDPLGRGNPLHLSRERYEKLHVMQHCIIIEIARSQFTNQ 1740
 1686 KARGCAYPAPYLDYGETDPLGRGNPLHLSRERYEKLHVMQHCIIIEIARSQFTNQ 1745
 1741 LFGFNWQLL 1749
 1726 LFGFNWQLL 1734

RESULT 7
 US-10-758-636A-19
 ; Sequence 19, Application US/10758636A
 ; Publication No. US20050089876A1

GENERAL INFORMATION:
 ; APPLICANT: Han, et al.
 ; TITLE OF INVENTION: HUMAN E3 ALPHA UBIQUITIN LIGASE FAMILY
 ; FILE REFERENCE: 01017/35966C
 ; CURRENT APPLICATION NUMBER: US/10/758,636A
 ; CURRENT FILING DATE: 2004-01-15
 ; PRIOR APPLICATION NUMBER: US 09/724,126
 ; PRIOR FILING DATE: 2000-11-28
 ; PRIOR APPLICATION NUMBER: US 60/187,911
 ; PRIOR FILING DATE: 2000-03-08
 ; NUMBER OF SEQ ID NOS: 29
 ; SOFTWARE: Patent in version 3.2
 ; SEQ ID NO 19
 ; LENGTH: 1734
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-758-636A-19

Query Match 98.5%; Score 9087; DB 17; Length 1734;
 Best Local Similarity 98.8%; Pred. No. 0;
 Matches 1728; Conservative 0; Mismatches 1; Indels 20; Gaps 1;

QY 1 MADBEAGGTERMEISABELPQTPQRLASWDDQVDFTYAFHLHLAQLVPEIYFAEMDPDLE 60
 DB 6 MADBEAGGTERMEISABELPQTPQRLASWDDQVDFTYAFHLHLAQLVPEIYFAEMDPDLE 65
 QY 61 KQESVQMSIPTLEWYLFGEDEPDICLEKLKHSAGFOLCGRVFKSGGETTYSRCDCAIDPT 120
 DB 66 KQESVQMSIPTLEWYLFGEDEPDICLEKLKHSAGFOLCGRVFKSGGETTYSRCDCAIDPT 125
 QY 121 CVLQWDCFDQSVHKNHRYKQHTSTGGGFCDCGDTTEAMWTKGPFVNHBPGRAGTIKENSRC 180
 DB 126 CVLQWDCFDQSVHKNHRYKQHTSTGGGFCDCGDTTEAMWTKGPFVNHBPGRAGTIKENSRC 185
 QY 181 PLNBEVIVQARKIPPSVIKYVVENTIWEBEKELPPELQIREKNERYCYVLFNDEHSHYDH 240
 DB 186 PLNBEVIVQARKIPPSVIKYVVENTIWEBEKELPPELQIREKNERYCYVLFNDEHSHYDH 245
 QY 241 VIYSLQALDCELAEALHTTAIDKEGRRAVKAGAYACQAEKEDIKSHSENVSOHPLHV 300
 DB 246 VIYSLQALDCELAEALHTTAIDKEGRRAVKAGAYACQAEKEDIKSHSENVSOHPLHV 305
 QY 301 EVLHSEIMAHQKPAALRLGSMNKMISYSSDFRQIFCQACLREBPDSNPCLISRLMLWDA 360
 DB 306 EVLHSEIMAHQKPAALRLGSMNKMISYSSDFRQIFCQACLREBPDSNPCLISRLMLWDA 365
 QY 361 KLYGARKILHELIPSSFPMEYKQLPAMEFVKYKQLOKEYISDDHDSISITALSVO 420
 DB 366 KLYGARKILHELIPSSFPMEYKQLPAMEFVKYKQLOKEYISDDHDSISITALSVO 425
 QY 421 MPTVPTLARHLIERQNVISVITETLLEVLPEYLDNRNKNFQGYSDQKLGKVYAVICDLK 480
 DB 426 MPTVPTLARHLIERQNVISVITETLLEVLPEYLDNRNKNFQGYSDQKLGKVYAVICDLK 485
 QY 481 YILISKPTIWTERRMQLEGFRSFLKILTCMQGMEERROVGOHIEVDPDWEAIAIQM 540
 DB 486 YILISKPTIWTERRMQLEGFRSFLKILTCMQGMEERROVGOHIEVDPDWEAIAIQM 545
 QY 541 QLNKILLMFQEWACDEBELLVAYKECHAVMRCTSPFISSKTVQSCGHSLETKSYRV 600
 DB 546 QLNKILLMFQEWACDEBELLVAYKECHAVMRCTSPFISSKTVQSCGHSLETKSYRV 605
 QY 601 SEDLVSIHLPLSRTLAGLHVRLSRGAVSRHLHEFVSPEDFQVEVLVEYPLRCLVLVAQVV 660
 DB 606 SEDLVSIHLPLSRTLAGLHVRLSRGAVSRHLHEFVSPEDFQVEVLVEYPLRCLVLVAQVV 665
 QY 661 AEMWERNGLSISQVFFYYQDVKCREMYDKDIIMLQIGASLMDPNKFLLLVLQRYELAE 720
 DB 666 AEMWERNGLSISQVFFYYQDVKCREMYDKDIIMLQIGASLMDPNKFLLLVLQRYELAE 725
 QY 721 FNKTIKTDQDLIKOYNTLIEEMQLVLIYIGERVVPGVGNVTKEETMBREIHLCLIEP 780
 DB 726 FNKTIKTDQDLIKOYNTLIEEMQLVLIYIGERVVPGVGNVTKEETMBREIHLCLIEP 785

Db 721 FNKTTSTKQDQLIKQYNTLIEBMLQVLIYVGERVYVGVGNVTRBEEVIMREITHLLCIEP 780
QY 781 MPHSAIAKVLPEVNNETGLENVINKVATPKPGVSGHGVYELKDSELDKFNMYFYHSK 840
Db 781 MPHSAIARLPEVNNETGLENVINKVATPKPGVSGHGVYELKDSELDKFNMYFYHSK 840
QY 841 TQHSKAEHMKQRRQKQENKDEALPPPPPPFCFAPSKVINLNCIDIMYILRTVPERAID 900
Db 841 TQHSKAEHMKQRRQKQENKDEALPPPPPPFCFAPSKVINLNCIDIMYILRTVPERAID 900
QY 901 TDSNLWTEGMLQWAFHILALGELLEKQLOKAPBEEVTFDFYHKASRLGSSAM---NTQM 957
Db 901 TESNLWTEGMLQWAFHILALGELLEKQLOKAPBEEVTFDFYHKASRLGSSAMNAQNIQM 960
QY 958 LLEKLGIPOLGEGOKDMITWILQMDFTVKRLREKSCLIIVATTSGSESINKDEITHDKKA 1017
Db 961 LLERLKGIPOLEGOKDMITWILQMDFTVKRLREKSCLIIVATTSGLECIKSEBITHDKEKA 1020
QY 1018 ERKRKAEARLHROKIMQMSALOKNPIETHKLMYDNTSEMPGKEDSIMEBESTPAVS DY 1077
Db 1021 ERKRKAEARLHROKIMQMSALOKNPIETHKLMYDNTSEVTKEDSIMEBESTSAVSEA 1080
QY 1078 SRIALGPKRGPSVTEKEVLTCTLCQEEQEVKIEENAMVLSACVQKSTALTQHRGKPIEL 1137
Db 1081 SRIALGPKRGPAVTEKEVLTCTLCQEEQEVKLENNAMVLSACVQKSTALTQHRGKPV D HL 1140
QY 1138 GBAIDPLFMDPLAVGTTTSGCHVHVAWCWKYFEAVQLSSQORIHVDLPDLEGEYLC 1197
Db 1141 GETDPLFMDPLAHGTTTSGCHVHVAWCWKYFEAVQLSSQORIHVDLPDLEGEYLC 1200
QY 1198 PLCKSICNTVPIPILOPKINSADALAQLLTLARWLOTVLARTSGVNIHAKGENP- 1256
Db 1201 PLCKSLCUNTPIPILOPKINSADALAQLLTLARWLOTVLARTSGVNIHAKGEAPA 1260
QY 1257 IPIFFNQMGDSTLEPHSILSGVSESSIKYSIKEMVILFATTIYRIGLKVPPDERDR 1316
Db 1261 VPVLFNQMGDSTLEPHSILSGVQSSVKYSIKEMVILFATTIYRIGLKVPPDELDPR 1320
QY 1317 VPMLTWSTCAFTIQAENLIGDEGKPLFGALQNRQNGKALKMOPAVAQRIICPVQLIOK 1376
Db 1321 VPMTWTSTCAFTIQAENLIGDEGKPLFGALQNRQNGKALKMOPAVAQRIICPVQLIHK 1380
QY 1377 HLVRLLSVVLPNLSKEDTCLLSIDLPHVLGAVLAPPSLYWDDPVDLPQSSVSSYNHL 1436
Db 1381 HLARLSVILPNLQSENTEGLLSVDLPHVLGAVLAPPSLYWDDTVDLPQSPSSYNHL 1440
QY 1437 YLFHLITMAHMLQILLTVDT---GLPLAQVQEDSEEAHSASSFFAIESQYTSIGCDI 1492
Db 1441 YLFHLITMAHMLQILLTDTDLSPGPPPLAEGEEDSEARCAFAFFVEVSQHTDGLTGCGA 1500
QY 1493 PGWYLVSLKNGITPYLRCAALPHYVLLGVTTPPELHTNSAGEYSALCSYLSLPTNPL 1552
Db 1501 PGWYLVSLRNGITPYLRCAALPHYVLLGVPPELFAHNSAGEFSALCSYLSLPTNPL 1560
QY 1553 LFOEYWDTRPLLQRCADPALLNCLQKNTVRYPRKNSLIELPDDYSCILLNOASHPR 1612
Db 1561 LFOEYWDTRPLLQRCWGGPALLKSLQKSAVVRYPRKNSLIELPEDYSCILLNOASHPR 1620
QY 1613 CPRSADDERKHPVLCFCGAILCSQNICCOEIVNGEVEGACIPFHALHCGAGVCIPLKIRE 1672
Db 1621 CPRSADDERKHPVLCFCGAILCSQNICCOEIVNGEVEGACVFPFHALHCGAGVCIPLKIRE 1680
QY 1673 CRVVLVEGARGCAYAPYLDYGETDPGLKRGNPPLHSRERYKRLHLVWQHCIIIEEIA 1732
Db 1681 CRVVLVEGARGCAYAPYLDYGETDPGLKRGNPPLHSRERYKRLHLVWQHCIIIEEIA 1740
QY 1733 RSQETNQMLPFGFNWQLL 1749
Db 1741 RSQETNQMLPFGFNWQLL 1757

RESULT 9

US-10-758-636A-15

; Sequence 15, Application US/10758636A
; Publication No. US20050089876A1
; GENERAL INFORMATION:
; APPLICANT: Han, et al.
; TITLE OF INVENTION: HUMAN E3 ALPHA UBIQUITIN LIGASE FAMILY
; FILE REFERENCE: 01017/35966C
; CURRENT APPLICATION NUMBER: US/10/758,636A
; CURRENT FILING DATE: 2004-01-15
; PRIOR APPLICATION NUMBER: US 09/724,126
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 60/187,911
; PRIOR FILING DATE: 2000-03-08
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 15
; LENGTH: 1757
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-758-636A-15

Query Match 93.4%; Score 8617; DB 17; Length 1757;
Best Local Similarity 92.1%; Pred. No. 0;
Matches 1619; Conservative 64; Mismatches 66; Indels 8; Gaps 3;
QY 1 MADEAGAGTTERMEIGABELPOTPPQRLASWMDQQVDFTAFLLHHLAQLVPEIYFAEMDPDLE 60
Db 1 MADEEMDGAERMDVSPPEPLAPQRPASWMDQQVDFTAFLLHHLAQLVPEIYFAEMDPDLE 60
QY 61 KOESVQMSIFTPLEWYLFGBDDPDI CLKLLKHSAGAFQICGRVFKSGGETTYSRCDCAIDPT 120
Db 61 KOESVQMSILTPLEWYLFGBDDPDI CLKLLKHSAGAFQICGRVFKSGGETTYSRCDCAIDPT 120
QY 121 CVLQMDCPQDSVHKHNRKYKMTSTGGGFCDCGTEAKTGPFCVNHPEGRAGTIIKENSRC 180
Db 121 CVLQMDCPQSSVHKHNRKYKMTSTGGGFCDCGTEAKTGPFCVNHPEGRAGTIIKESLHC 180
QY 181 PLNREEVIVQARKIPPSVIKYVEMTIWBEKEKLPPELQIREKNERYCYVLFNDEHSHSYDH 240
Db 181 PLNREEVIAQARRIFPSVIKYVEMTIWBEKEKLPPELQIREKNERYCYVLFNDEHSHSYDH 240
QY 241 VYISLQALDCLABRAQLHTTAIDKEGRRAVKAGAYACQEAKEDEIKHSENVSHQPLHV 300
Db 241 VYISLQALDCLABRAQLHTTAIDKEGRRAVKAGAYATCQEAKEDEIKHSENVSHQPLHV 300
QY 301 EYLHSEIWAHQFALRLGSMWNKINSYSSDFQIIFCOACLRPEPDSNPCLISRLMLWDA 360
Db 301 EYLHSEVWMAHQFALRLGSMWNKINSYSSDFQIIFCOACLRPEPDSNPCLISRLMLWDA 360
QY 361 KLYKGARKILHELIFSSFFMEMEYKCLFAMEFVKYKQLOKEYISDDHRSISITALSVQ 420
Db 361 KLYKGARKILHELIFSSFFMEMEYKCLFAMEFVKYKQLOKEYISDDHRSISITALSVQ 420
QY 421 MFTVPTLARHLIEBQNVISVITETLLEVLPEYLDNRNKNFQGYSDQKLGRTVAVICDLK 480
Db 421 MLTVPTLARHLIEBQNVISVITETLLEVLPEYLDNRNKNFQGYSDQKLGRTVAVICDLK 480
QY 481 YLILSKPTIWTRELRAQFLBGRSFLKILTCMQGMEIRROVGOHLEVDPDMEAAIAIOM 540
Db 481 YLILSKPTIWTRELRAQFLBGRSFLKILTCMQGMEIRROVGOHLEVDPDMEAAIAIOM 540
QY 541 QLNKILLMPQEWACDEDELLLVAYKECHKAVNRCSSTFSSSKTVVQSCGHSLETYSYV 600
Db 541 QLNKILLMPQEWACDEDELLLVAYKECHKAVNRCSSTFSSSKTVVQSCGHSLETYSYV 600
QY 601 SEDLVSIHLPLSRTLAGLHVLSRLGAVSRHLHEFVSFEDFQVEVLVEYPLCLVLVAQV 660
Db 601 SEDLVSIHLPLSRTLAGLHVLSRLGAVSRHLHEFVSFEDFQVEVLVEYPLCLVLVAQV 660
QY 661 AEWWRNGLSLSQVFTYODVKCRBEMDKDIIIMLQIGASLMDPNKFLLLVLQRYELAE 720
Db 661 AEWWRNGLSLSQVFTYODVKCRBEMDKDIIIMLQIGASLMDPNKFLLLVLQRYELTDA 720
QY 721 FNKTTSTKQDQLIKQYNTLIEBMLQVLIYVGERVYVGVGNVTRBEEVIMREITHLLCIEP 780

Db 721 FNKTIISKODLLIKQNTLIEEMQLVLIYVGRYVPGVGNVTRREVMRBIHLJCIEP 780
Qy 781 MPHSALAKNLPENNETTGLENVINKVATPKPGVSGHGYELKDBSLKDFNMFYHYSK 840
Db 781 MPHSALARNLPENNETTGLENVINKVATPKPGVSGHGYELKDBSLKDFNMFYHYSK 840
Qy 841 TQSKAHEHQKRRKQENKDEALPPPPPPPPCFAPSKVINLNCDDMMVILRTVFRRAID 900
Db 841 TQSKAHEHQKRRKQENKDEALPPPPPPPPCFAPSKVINLNCDDMMVILRTVFRRAID 900
Qy 901 TDSNLTGEMQWAFHILALGLLEEKQOLQKAPBEVEVTFDVFHKA SRLGSSAM--NIOM 957
Db 901 TDSNLTGEMQWAFHILALGLLEEKQOLQKAPBEVEVTFDVFHKA SRLGSSAMNIOM 960
Qy 958 LLEKLAGIPQLEGQKDMITWILQMDFTVKRLREKSLIVATTSGSISIKNDIETHDKKA 1017
Db 961 LLEKLAGIPQLEGQKDMITWILQMDFTVKRLREKSLIVATTSGSISIKNDIETHDKKA 1020
Qy 1018 ERKKAARLHRQKIMAKSALQKNPIETHKLMYDNTSEMPKCKDSIMEESTPAVS DY 1077
Db 1021 ERKKAARLHRQKIMAKSALQKNPIETHKLMYDNTSEMPKCKDSIMEESTPAVS DY 1080
Qy 1078 SRIALGPKRGPSTKEVLTCLCOEQQVKIENNAWLSACVOKSTALTQHRGKPIELS 1137
Db 1081 SRIALGPKRGPSTKEVLTCLCOEQQVKIENNAWLSACVOKSTALTQHRGKPIELS 1140
Qy 1138 GBALDPLFMDPOLAYTYTSGCHVNHAVCWQYFEAVOLSSOORTHVDLFDLESSEYLC 1197
Db 1141 GETDPLFMDPOLAYTYTSGCHVNHAVCWQYFEAVOLSSOORTHVDLFDLESSEYLC 1200
Qy 1198 PLCKSLCNTVPIPILOPKINSENADALQTLTARWITQTLARISGVNIRHAKGNP- 1256
Db 1201 PLCKSLCNTVPIPILOPKINSENADALQTLTARWITQTLARISGVNIRHAKGNP- 1260
Qy 1257 IPFPNOGQDSTLPHSILSGVSSIKYNSIKEMVILPATTIYRIGLKVPPDRDR 1316
Db 1261 IPFPNOGQDSTLPHSILSGVSSIKYNSIKEMVILPATTIYRIGLKVPPDRDR 1320
Qy 1317 VPMLTWSTCAFTIQTALENLIDBEGKPLFGALQNRQNGHKKALMOPAVOARITCPOVLIOK 1376
Db 1321 VPMWTWSTCAFTIQTALENLIDBEGKPLFGALQNRQNGHKKALMOPAVOARITCPOVLIOK 1380
Qy 1377 HLVRLLSVVLPNIKSDPTCLLSIDLPHVLGAVLAFPSLYWDDPVDLPQSSVSSYNHL 1436
Db 1381 HLARLLSVLPLNIOSENTTGLLSVDLPHVLGAVLAFPSLYWDDPVDLPQSSVSSYNHL 1440
Qy 1437 YLFHLITWAHMLQILLTVDTTDLSPGPPPLAEGBEDSEARCAFAFVVSQHTDGLTGCGA 1500
Db 1441 YLFHLITWAHMLQILLTVDTTDLSPGPPPLAEGBEDSEARCAFAFVVSQHTDGLTGCGA 1500
Qy 1493 PGWYLVWSLNGITPVLRCALPHYLIGVTPPELHTNSARGEYSALCSYLSLPTNLP 1552
Db 1501 PGWYLVWSLNGITPVLRCALPHYLIGVTPPELHTNSARGEYSALCSYLSLPTNLP 1560
Qy 1553 LFQEWYDVTVPRLQWCAADPALLNCLIKQNTVVPYPRKNSILIELPDDYSCLLNOASHPR 1612
Db 1561 LFQEWYDVTVPRLQWCAADPALLNCLIKQNTVVPYPRKNSILIELPDDYSCLLNOASHPR 1620
Qy 1613 CPRSADDERKHPVLCFCGAILCSQNICCOEIVNGEEVGACITPHALHCGAGVCIPLKIRE 1672
Db 1621 CPRSADDERKHPVLCFCGAILCSQNICCOEIVNGEEVGACITPHALHCGAGVCIPLKIRE 1680
Qy 1673 CRVVLVEGKARGCAPAPYLDYGETDPLGKGNPLHLSREYRKHLVWQOHCIIIEBIA 1732
Db 1681 CRVVLVEGKARGCAPAPYLDYGETDPLGKGNPLHLSREYRKHLVWQOHCIIIEBIA 1740
Qy 1733 RSQETNQMLFGFNWQLL 1749
Db 1741 RSQETNQMLFGFNWQLL 1757

US-10-758-672A-4
; Sequence 4, Application US/10758672A
; Publication No. US20040185037A1
; GENERAL INFORMATION:
; APPLICANT: Han, et al.
; TITLE OF INVENTION: HUMAN E3 ALPHA UBIQUITIN LIGASE FAMILY
; FILE REFERENCE: 01017/35966B
; CURRENT APPLICATION NUMBER: US/10/758,672A
; PRIOR FILING DATE: 2004-01-15
; PRIOR APPLICATION NUMBER: US 09/724,126
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 60/187,911
; PRIOR FILING DATE: 2000-03-08
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 4
; LENGTH: 1755
; TYPE: PRT
; ORGANISM: Homo sapeins
US-10-758-672A-4

Query Match 46.9%; Score 4327; DB 16; Length 1755;
Best Local Similarity 47.2%; Pred. No. 0;
Matches 841; Conservative 339; Mismatches 543; Indels 58; Gaps 27;

Qy 1 MADE---EAGGTGTER--MEISAEPLQTPQRLASWMDQVDFYTAFLHHLAQLVPEIYFAEM 55
Db 1 MASELEPEVQADRSILCEAS-----EIAKMLQATDLTREVYQHLAHHVYPKIYCRGP 54
Qy 56 DPDLSEKQESVQMSPTPLEWYLFQEDDPDICLEKHKISG-AFQLCGRVPKSGSETTYSCRD 114
Db 55 NFPPOKEDMLAQLHVLGPMWYLCGEDPAFGPGLKLEQANKPSHLCKGRVFKVGEPTYSCRD 114
Qy 115 CAIDPTCVLCMDQFQDSVHKHRYKMHSTGSGFCDCGDEATKGTGPPCVNHEPGRAG-T 173
Db 115 CAVDPTCVLCMECFGSHRDHRYMTTSGGGGFCDCGDEATKGTGPPCVNHEPGRAG-T 174
Qy 174 IKENSRCPLNEVIVQARKIPFVIVVEMTWEEKELPPELOIREKNERYCVLFFND 233
Db 175 BEEDPLVHLSVEDVIARTYNIPATITRYAVBIILTWEKSELPADELVKESDYTCMLFND 234
Qy 234 EHSYDHYVYSIQRALDCELAQAHLHTTAIDKEGRRAVAGAYAAQCEAKEDIKSHSENV 293
Db 235 EVHTYEQVITYTLQKAVNCTQKEAIGFATTVDRGRSVRYGDFQYCEQAKSVIVRNTSRQ 294
Qy 294 SOHPLHVEVLHSEIMAHQKFAIRLGLSNKMSYSSDPRIQIFCOACLREPPSENPLIS 353
Db 295 TK-PLKQVMHSSIVAHQNFGLKLSWLGSIIIGSDGLRRILCQVGLQSGPGGNSSLVD 353
Qy 354 RLMLMDAKLYKGAARKILHELIPSSPFMEBYKCLFAMBFVKYKQLOKEIYISDDHDSRIS 413
Db 354 RLMLSDSKLWKGARSVYHQLFMSSLLMDLKYKCLFAVRFAKQYQQLORDPMEDDHERAVS 413
Qy 414 ITALSVMQVTPVTLARHLIRBONVIVTETLLEVLPEYLDNRNKNFNGYS---ODKLG 470
Db 414 VTALSVPQFTAPTARMLITEENLSIIKTFMDHL-RHRDAQGRFQFERYTALQAFKER 472
Qy 471 RVYAVICDILKYLLISKPTIWTERRLMQFLEGRSPFLKILTCHQGMEEIRRVQGHIEVDP 530
Db 473 RVQSLILDKVLYLSKPTTENSDELKQFLQEGDAFLELLKCMQGMDDPITRVQGHIEVDP 532
Qy 531 DWEAAIAIQMOLKNILLMFOEWCACDEBELLVAYKSKCHAVMRCSTSFSSSKTVQS-C 589
Db 533 EWEAAFTQMKLTHVISMQDWCADEKVLIEAYKKCLAVLMQCHGGYTDGQPTILSLIC 592
Qy 590 GHSLETSKYRSVEDLSVSIHLPLSRTLAGLVRSLGAVSRHLHFEVSPEDFOVFLVEYVP 649
Db 593 GHSVETIRYCVSQEKVSIHLPSVRLLAGLHVLSSKSEVAYFPPELLPLSELPPMLIEHP 652
Qy 650 LRCLVLVAQVAVMWRNGLSLISQVFTYQDVCKREMYDIDIMLOIGASLMDNPKPL 709
Db 653 LRCLVLCQAQVHAGWRRNGPSLVNQIYTYHNKCRREMFDKDVVLMQTVGVSMDDPNHFLM 712

Qy	1126	LTQHRKPIELSGEALDPLFMDPOLAYGTGTCGHWMAHVCWQKTFEAVQLSSQ----	1181
Db	656	LSKRSKFIO-DPEKDYDPLFPHFDLSCGTHSSCGHMHACWQRFYDYSVQAKBORRQOR	714
Qy	1182	-RIHVDLFDLSEGEYLCPLCKSLCNTVPIPIPOQKINSADAAQLLTLTARWTQTVL	1240
Db	715	LRLHTS-YDVENGELCPLCECLSNITVPLL-LPPRNIFNNRLN-PSDOPNLQTOWIRITS	771
Qy	1241	ARISGVNIRHAKGENPIPIFNQMGDSTLEFHSILLSFGVSESSIKYSNKEMVILFATT	1300
Db	772	QQIKALQFLRKESSTPNASTKNSENVDEIQLEPGFRPDRPKPIPYSESIKEMLTTFGTA	831
Qy	1301	IYRIGLKVPDPDRPRVPMLTWTSCAFTIOAIENLLGDEGKPLFGALQNRQHNGLQALMQ	1360
Db	832	TYKVLKVHNEEDPRVIMCWGSCAVTIQSIERILSDCKPLFGPLPCRLDDCLASLNR	891
Qy	1361	FAVAQRITCPQVLIQKHLVRLLSVVLPNIKSEBTPCLLSIDLPHVLGVAVLAPPSLYWDD	1420
Db	892	FAAAHWTVASVSVQGHCFKLFASLVENDSHEELPCILDDMFHLLVGLVLAFFALPQCD	951
Qy	1421	PVDLQPSVSSVNNHLYLPHLITMAHMLQIL---TVDTCGLPAQVEDSEEAHSASSPF	1477
Db	952	-----PSGISLGTGDHLHFLHVTMAHIIQLLTSCTEENGMD--QENPPCEEBSAVALY	1004
Qy	1478	ABISOYTSGSIGCDIP--GWYLWLSLKNGITPYLRCAALPFHYLLGVTTPPBLHTNSAEGE	1536
Db	1005	KTLHQYT-GSALKKEIPSGMHLWRSVRAGIMPFLKCSALFFHYLNGVSPSPDIQV-PGTSH	1062
Qy	1537	YSALCSYLSLPTNLFLLPQBYWDTVRPLQRCWADPALMCLKOKNTVVRYPRKRSNLIE	1596
Db	1063	FEHLCSYLSLPPNNLICLFQENSEIMNLSLISWCNKEVKRYLEGERDAIRYPRESNKLIN	1122
Qy	1597	LPDYSCLLNQASHFRCPERADDERKHPVLICFGAILCSQNTCCQEIYNGVEEGVACIEH	1656
Db	1123	LPEDYSLINQASNFSPKSGGDKSRAPTIICLVGSLLSQSCCOTELGEEDVGACTAH	1182
Qy	1657	ALHCGAGVCYFLKIKREKRVVLVEKGARGCAYPAPYLDVEYGETOPLKRGNPMLHLSRERY	1716
Db	1183	TYSCSGVGWFLRVREQVFLAGTKGCFYSPPYLDDYGETDQGLRGNPLHLCKERPK	1242
Qy	1717	KHLVWQOQHCHIEBIARSQETNQMQLGFQNWQLL	1749
Db	1243	KIOKIWHOSHVSYTERIGHAORANOTLVGIDWOHL	1275

RESIT.T 15

RESULTS
US-10-094-749-2555

US-10-094-749-2333 : Sequence 2555. Application US/10094749

; sequence 2333, Application US/1
: Publication No. US20030219741A1

; PUBLICATION NO: US20
: GENERAL INFORMATION:
: GENERAL INFORMATION:

; GENERAL INFORMATION: TSOGAT TAKAO

APPLICANT: ISOGAI, TAKAO
APPLICANT: SUGIYAMA TOMOYASU

APPLICANT: SUGIYAMA, TOMOYUKI
APPLICANT: OTSUKI, TETSUHIRO

APPLICANT: OTSUKI, TETSUJI
APPLICANT: WAKAMOTO, AT

APPLICANT: WAKAMATSU, AI
 REPORT DATE: 11 NOV 67

; APPLICANT: SATO, HIROYUKI

APPLICANT: ISHII, SHIZUKO

APPLICANT: YAMAMOTO, JUN-1

APPLICANT: ISONO, YUUKO

APPLICANT: HIO, YURI

APPLICANT: OTSUKA, KAORU

APPLICANT: NAGAI, KEIICHI

APPLICANT: IRIE, RYOTARO

APPLICANT: TAMECHIKA, ICHIRO

APPLICANT: SEKI, NAOHIKO

APPLICANT: YOSHIKAWA, TSUTSUMI

; APPLICANT: OTSUKA, MOTOYUKI

APPLICANT: NAGAHARI, KENJI

APPLICANT: MASUHO, YASUHIKO

1. TITLE OF INVENTION: NOVEL FU

: FILE REFERENCE: 084335/0160

FILE REFERENCE: 084335/0100
CURRENT APPLICATION NUMBER: US/10

```

; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2555
; LENGTH: 594
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-2555

```

Query Match 33.2%: score 3066: DB 15: Length 594:

Query Match 33.2%, Score 3000, DB 157,
Best [local] similarity 100.0%: Pred. No. 2-8e-243:

Best Local Similarity 100.0%; Pred. NO: 2.8e-243;
Matches 594: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Qy	421	MFTVPTLARHLIEBQNVISVITETLLEVLPEYLDNRNKNFQGSODKLG RVYAVICDLK	480
Db	1	MFTVPTLARHLIEBQNVISVITETLLEVLPEYLDNRNKNFQGSODKLG RVYAVICDLK	60
Qy	481	YILISKPTIWTERRMOPLEGFRSLKILTCMQGMEEIRROVQGHIEVDPDWEAAIAIQM	540
Db	61	YILISKPTIWTERRMOPLEGFRSLKILTCMQGMEEIRROVQGHIEVDPDWEAAIAIQM	120
Qy	541	QLKNILMFOQWCADEBELLVAYKECHAVMRCTSTFISSKTVVQSGCHSLETKSYRV	600
Db	121	QLKNILMFOQWCADEBELLVAYKECHAVMRCTSTFISSKTVVQSGCHSLETKSYRV	180
Qy	601	SEDIVSIHLPISRTLAGLHVRLSRLGAVSRLLHFVSPEDQVEVLVVEYPLRCLVLVAQVV	560
Db	181	SEDIVSIHLPISRTLAGLHVRLSRLGAVSRLLHFVSPEDQVEVLVVEYPLRCLVLVAQVV	240
Qy	561	AEWRRNGLSLISQVFFYQDVKCBREMYDKDIIMLOIGASLMDPNKFLLLVLQRYELAEA	720
Db	241	AEWRRNGLSLISQVFFYQDVKCBREMYDKDIIMLOIGASLMDPNKFLLLVLQRYELAEA	300
Qy	721	FNKTIISTKQDLIIKOYNTLIEBMLQVLIYIYIGERVPGVGNVTKEEVTMRREIHLLCIEP	780
Db	301	FNKTIISTKQDLIIKOYNTLIEBMLQVLIYIYIGERVPGVGNVTKEEVTMRREIHLLCIEP	360
Qy	781	MPHSAIAKNLPENENNETGLENVINKVATPKPGVSGHGVBELKDESXKDPNMYFHYYSK	840
Db	361	MPHSAIAKNLPENENNETGLENVINKVATPKPGVSGHGVBELKDESXKDPNMYFHYYSK	420
Qy	841	TOHSAEAMQKRRKOENKDEALPPPPPEPCPAFSKVINLNCDIMMYILRTVFFERAID	900
Db	421	TOHSAEAMQKRRKOENKDEALPPPPPEPCPAFSKVINLNCDIMMYILRTVFFERAID	480
Qy	901	TDSNLTWTEGMLQMAFHIIALGLLEBKQOLQKAPEBEVTFDFVHKASRLGSSAMNIOMLLE	960
Db	481	TDSNLTWTEGMLQMAFHIIALGLLEBKQOLQKAPEBEVTFDFVHKASRLGSSAMNIOMLLE	540
Qy	961	KLKGIPQLEGOXDMITWILQMPDVTVKRLREKSLIVATTSGSESINKDEITHDK	1014
Db	541	VLKGBOLREGOKMTWILQMPDVTVKRLREKSLIVATTSGSESINKDEITHDK	594

Search completed: July 9, 2005, 15:43:25

Job time : 142 sec

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 9, 2005, 13:43:17 ; Search time 47 Seconds
(without alignments)
3580.492 Million cell updates/sec

Title: US-10-758-672A-2
Perfect score: 9224
Sequence: 1 MADEEAGGTERMEISAEPLQ.....ETARSQETNQLGFNWL 1749

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8617	93.4	1757	2 T14318	ubiquitin-protein
2	1945.5	21.1	1927	2 T25604	hypothetical prote
3	1038	11.3	1958	2 T39808	hypothetical prote
4	914	9.9	2052	2 T37711	probable n-end-rec
5	689.5	7.5	1950	2 S12332	ubiquitin-protein
6	647.5	7.0	1941	2 T30554	ubiquitin-protein
7	606.5	6.6	1225	2 T48251	ubiquitin-protein
8	549.5	6.0	1872	2 S64851	probable membrane
9	537.5	5.8	2168	2 D88131	protein F10G7.10 (
10	343	3.7	795	2 T48252	eceriferum3 (CERS)
11	218	2.4	271	2 T40238	hypothetical prote
12	205.5	2.2	3187	2 JCS5837	364K Golgi complex
13	190.5	2.1	3259	1 A56539	giantin - human
14	186	2.0	1410	1 A57013	early endosome ant
15	182	2.0	2442	2 T08621	centrosome associa
16	180.5	2.0	1413	2 T28467	hypothetical prote
17	180	2.0	1830	1 S19188	myosin-V - chicken
18	179.5	1.9	2346	2 T13829	Tpr homolog - frui
19	176	1.9	3225	2 I52300	giantin - human
20	175.5	1.9	1427	2 S22695	restin - human
21	175.5	1.9	1526	2 T41522	myosin ii - fission
22	175.5	1.9	2863	1 S28261	centromere protein
23	174.5	1.9	1009	2 C89910	hypothetical prote
24	174.5	1.9	1392	2 A43336	microtubule-vesicl
25	172.5	1.9	1780	2 T17272	hypothetical prote
26	168	1.8	1837	2 T41023	probable nuclear p
27	166	1.8	1163	2 G97236	ATPase involved in
28	166	1.8	1676	2 E71410	probable centromer
29	164.5	1.8	1181	2 T01799	hypothetical prote

30	164	1.8	2954	2 T14156	kinesin-related pr
31	163.5	1.8	1339	2 H81307	restriction modifi
32	163	1.8	1853	1 A46761	myosin heavy chain
33	163	1.8	2649	2 A40937	bullous pemphigoid
34	162	1.8	1269	2 F84730	probable myosin he
35	157.5	1.7	1642	2 T08880	NMDA receptor-bind
36	156.5	1.7	1496	2 T05634	hypothetical prote
37	156.5	1.7	1790	2 S67593	transport protein
38	156.5	1.7	2710	2 A37052	reticulocyte-bindi
39	156.5	1.7	2829	2 A42771	toxin A - Clostrid
40	155.5	1.7	1216	2 A28821	1-phosphatidylinos
41	155	1.7	1025	2 S86355	hypothetical prote
42	155	1.7	1173	1 A53430	1-phosphatidylinos
43	155	1.7	1828	2 B59254	myosin heavy chain
44	154.5	1.7	2269	2 T28677	rhopty protein -
45	154	1.7	3685	1 A27605	dystrophin, muscle

ALIGNMENTS

RESULT 1

T14318
ubiquitin-protein ligase E3-alpha - mouse
N:Alternate names: N-recognin E3-alpha
C:Species: Mus musculus (house mouse)
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
R:Kwon, Y.T.; Reiss, Y.; Fried, V.A.; Herahko, A.; Yoon, J.K.; Gonda, D.K.; Sanghan, P.;
Proc. Natl. Acad. Sci. U.S.A. 95, 7898-7903, 1998
A:Title: The mouse and human genes encoding the recognition component of the N-end rule
A:Reference number: Z17977; MUID:98318583; PMID:9653112
A:Accession: T14318
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1757 <KWO>
A:Cross-references: UNIPROT:O70481; EMBL:AF061555; MID:g3170886; PID:g3170887; PIDN:AAC
C:Genetics:
A:Gene: Ubr1
A:Map position: 2

Query Match	93.4%	Score	8617	DB	2	Length	1757
Best Local Similarity	92.1%	Pred. No.	0				
Matches	1619	Conservative	64	Mismatches	66	Indels	8
Gaps	3						
QY	1	MAD EAGGTERMEISAEPLQTPQRLASWWDQVDFTAFLLHQAQLVPEIYFAEMDPDL	60				
DB	1	MAD EMDGAEKMDVSPPEPLAPQRPASWWDQVDFTAFLLHQAQLVPEIYFAEMDPDL	60				
QY	61	KQESVQMSIPTLEWYLFGEEDPDICLEKLKHSQAFQICGRVFKSGETTYSCRDCAIDPT	120				
DB	61	KQESVQMSILTPLEWYLFGEEDPDICLEKLKHSQAFQICGRVFKSGETTYSCRDCAIDPT	120				
QY	121	CVLCHDCQDSVHKHRYKMTSTGGCDGDTAWKTPPCVNHFPGRAGTIKENSRC	180				
DB	121	CVLCHDCQSSVHKHRYKMTSTGGCDGDTAWKTPPCVDHDFGRAGTTKESLHC	180				
QY	181	PLNEEVIQAARKIPPSVIKYVEMTWEEKELPELOIRREKNERYICVLPNDEHSHVDH	240				
DB	181	PLNEEVIQAARKIPPSVIKYVEMTWEEKELPELOIRREKNERYICVLPNDEHSHVDH	240				
QY	241	VIYSLQALDCELAQAQLHTTAIDKGGRAVKAGAYACQBAKEDIKSHSENVSOHPLHV	300				
DB	241	VIYSLQALDCELAQAQLHTTAIDKGGRAVKAGAYATCQBAKEDIKSHSENVSOHPLHV	300				
QY	301	EVLHSEIHAHQFALRLGSSWANKMSYSSDFRQIFCOACLBEEPDPSENPCLISRLMLWA	360				
DB	301	EVLHSEIHAHQFALRLGSSWANKMSYSSDFRQIFCOACLBEEPDPSENPCLISRLMLWA	360				
QY	361	KLYGARKILHELIFSSPFMEYKCLFAMFVKYKQLOKEYISDDHDSISITALSVDQ	420				
DB	361	KLYGARKILHELIFSSPFMEYKCLFAMFVKYKQLOKEYISDDHDSISITALSVDQ	420				

QY 421 MFTVPTLARHLIEQNVIIVITETILLVPEYLDRNNKENFOGYSQDKLGRVAVICDLK 480
DB 421 MFTVPTLARHLIEQNVIIVITETILLVPEYLDRNNKENFOGYSQDKLGRVAVICDLK 480
QY 481 YILISKPTIWTBRLRMQFLEGRSFLKILTCMQGMBEIRROVGOHIEVDPDWEAAIAIQM 540
DB 481 YILISKPTIWTBRLRMQFLEGRSFLKILTCMQGMBEIRROVGOHIEVDPDWEAAIAIQM 540
QY 541 QLNKILLMFOEWCACDELLLVAYKECHAVNRCSSTPSSKTVVQSCGHSLETKSYK 600
DB 541 QLNKILLMFOEWCACDELLLVAYKECHAVNRCSSTPSSKTVVQSCGHSLETKSYK 600
QY 601 SEDLSIHLPLSRTLAGHVRSLGAVSRLEHFVSFEDPOEVLVPEYPLCLVLVAQVV 660
DB 601 SEDLSIHLPLSRTLAGHVRSLGAVSRLEHFVSFEDPOEVLVPEYPLCLVLVAQVV 660
QY 661 AEMWRREGLSLISQVPIYQDVKREEMDKDILMLQIGASLMDPNKFLLLVLQRYELAEA 720
DB 661 AEMWRREGLSLISQVPIYQDVKREEMDKDILMLQIGASLMDPNKFLLLVLQRYELAEA 720
QY 721 FNKTISTKQODLIKQNTLIEBMLQVLIYVGERVYVGVNTKBEVTVREIILLICIRP 780
DB 721 FNKTISTKQODLIKQNTLIEBMLQVLIYVGERVYVGVNTKBEVTVREIILLICIRP 780
QY 781 MPHSATAKNLPENNNETGLENVINKVATFKPGVSGHGVYELKDBSLKDFNMFVYHSK 840
DB 781 MPHSATAKNLPENNNETGLENVINKVATFKPGVSGHGVYELKDBSLKDFNMFVYHSK 840
QY 841 TQHSKAEHMKQKRKQENKDEALPPPPPEFCFAPSKVINLNCIDIMVILRTVFERAID 900
DB 841 TQHSKAEHMKQKRKQENKDEALPPPPPEFCFAPSKVINLNCIDIMVILRTVFERAID 900
QY 901 TQSNLWTEGMLQWAFHILALGLEEKQLOKAPEREVTFDFYHKASRLGSSAM---NIQM 957
DB 901 TQSNLWTEGMLQWAFHILALGLEEKQLOKAPEREVTFDFYHKASRLGSSAMNAQIQM 960
QY 958 LLEKLGIPQLEGQKDMITWILQMPDFTVRLREKSLIVATTSGSBSIKNDRIETHDKKA 1017
DB 958 LLEKLGIPQLEGQKDMITWILQMPDFTVRLREKSLIVATTSGSBSIKNDRIETHDKKA 1020
QY 1018 ERKKAABARHLRQKIMQMSALQKNFIETHKLYMDNTSEMPGKEDSIMEESTPAVSIDY 1077
DB 1021 ERKKAABARHLRQKIMQMSALQKNFIETHKLYMDNTSEVTKGKEDSIMEESTSAVSEA 1080
QY 1078 SRIALGPKRGPSVTEKVLTCILCOBEQEVKIENNAMVLSACVQKSTALTQHRGKPIELS 1137
DB 1081 SRIALGPKRGPAVTEKVLTCILCOBEQEVKIENNAMVLSACVQKSTALTQHRGKPEVDHL 1140
QY 1138 GEALDPLFMDPDLAYGTYTSCGCHVMHVCWKYFEAVQLSSOORIHVDLPDLESGEYLC 1197
DB 1141 GETLDPLFMDPDLAHGTYTSCGCHVMHVCWKYFEAVQLSSOORIHVDLPDLESGEYLC 1200
QY 1198 PLCKSLCNTVPIPIPLQPOKINSENADALQALLTLARWITQVLARISGVNIRHAKGENP- 1256
DB 1201 PLCKSLCNTVPIPIPLQPOKINSENADALQALLTLARWITQVLARISGVNIRHAKGEAPA 1260
QY 1257 IPIFFNQMGDSSTLFPHSILSPGVSSIKYSNIKEMVILFATTIYRIGLKVPPDERDR 1316
DB 1261 VPVLFNQMGDSSTLFPHSILSPGVSSIKYSNIKEMVILFATTIYRIGLKVPPDELDPR 1320
QY 1317 VPMLTSTCAFTIQAIENTLGDGKPLFGALQNRQHGLKALMQFAVAORATPCQVLIQK 1376
DB 1321 VPMMTWTSTCAFTIQAIENTLGDGKPLFGALQNRQHGLKALMQFAVAORATPCQVLIHK 1380
QY 1377 HLVRLLSVLPNTKSEDTPLCLSIDLPHVLVGLVAFPPSLYMDPDVLOPSSVSSSYNHL 1436
DB 1381 HLARLLSVLPNLOSENTGELLSDVDFPHVLVGLVAFPPSLYMDPDVLOPSSVSSSYNHL 1440
QY 1437 YLFHLITMAHMLQILLTVDT-----GLPLAQVQEDSEAHASAFPFAEISQYTSGSTGCDI 1492
DB 1441 YLFHLITMAHMLQILLTVDT-----GLPLAQVQEDSEAHASAFPFAEISQYTSGSTGCGCA 1500
QY 1493 PGWYLWVSLKNGITPYLRCAALFFHYLLGVTPPEELHTNSAEYSALCSYLSLPTNLFL 1552

DB 1501 PGWYLWVSLKNGITPYLRCAALFFHYLLGVAPPRELPAASAEGEFPSCSYLSLPTNLFL 1560
QY 1553 LFOEYWDVTBRLQWCADPALLNCLKOKNTVTVPRGRNSLIELPDDYSCLLNQASHPR 1612
DB 1561 LFOEYWDVTBRLQWCGDPALLKSLKQKSAVVYPRKNSLIELPDDYSCLLNQASHPR 1620
QY 1613 CPBSADDERKHPVLCFLCGAILCSQNICCOBTVNGEEVGACIFHALHCGAGVCIFFLKIRE 1672
DB 1621 CPBSADDERKHPVLCFLCGAILCSQNICCOBTVNGEEVGACIFHALHCGAGVCIFFLKIRE 1680
QY 1673 CRVVLVEGKARGCAYPAPYLDVEYGETDPGLKRGNEPLHLSRERYRKLHLVWQOHCIEETA 1732
DB 1681 CRVVLVEGKARGCAYPAPYLDVEYGETDPGLKRGNEPLHLSRERYRKLHLVWQOHCIEETA 1740
QY 1733 RSQETNQMLFGFNWQLL 1749
DB 1741 RSQETNQMLFGFNWQLL 1757

RESULT 2
T25604
hypothetical protein C3288.11 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T25604
R:Gattung, S.
submitted to the EMBL Data Library, February 1997
A:Description: The sequence of C. elegans cosmid C3288.
A:Reference number: Z20056
A:Accession: T25604
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1927 <GAT>
A:Cross-references: UNIPROT:P91133; EMBL:U88308; PIDN:AAB42328.1; GSPDB:GN00019; CESP:C3
A:Experimental source: strain Bristol N2; clone C3288
C:Genetics:
A:Gene: CESP:C3288.11
A:Map position: 1
A:Introns: 56/3; 157/3; 255/3; 469/1; 578/3; 832/3; 907/2; 1165/2; 1218/2; 1275/2; 1468/

Query Match 21.1%; Score 1945.5; DB 2; Length 1927;
Best Local Similarity 27.6%; Pred. No. 2.5e-111;
Matches 560; Conservative 311; Mismatches 663; Indels 493; Gaps 66;

QY 97 QLCGRVFSGETTYSRCDCAIDPTCVLCMDQFQDSVHKHRYMYMTSTGGGFCDCGDTGA 156
DB 14 QICGHVFKNGLTYTCLDCATDGTCTVMCLQCPEVSIHKSHKYMHSMSGSGYCDGCDADA 73
QY 157 WKTGPPCVNHEPG-----RAGTIKENSRCPLNEEVIVQARKIPPSVLYKYVEMTWEE 209
DB 74 WTEGYACANHEKKODEEAAVLAPELKK--RCSQLVEIILQ---FSLSMITHKDDLLKLP 128
QY 210 EKELPELQIREKNERYCVLFENDEHSHSDYHVIYSIQRALDCELAEALHTTAIDKEGRR 269
DB 129 FEKMEPE--VTNEAQYLTLYNDETHYVESIVKLELYIHTCTQQAAMLVATVDREGS 186
QY 270 AVKAGAYAACQBAKEDIKSHS-----ENVSOH--PLHVEVLHSEIMAHQKPAKRLGWS 320
DB 187 AVKLGSKADCTKADKDDVQKRTADPTSIKSSNNHNLPLSVKVMDDTTLFALQNFSLIATW 246
QY 321 MNKIMSYSDFFQIQCOACL----- 340
DB 247 LNTQMDVFPPLREIVGEILLSSKFKLKNYTRMKSEDOQLVAGIIRNVNVLPPDEEEL 306
QY 341 -----REEP-----DSENPCLI 352
DB 307 PALDGRMDVEMDDDDIGCALQMEIDADDEBEITAAALAGVSEHQSSPGSRDSTFTML 366
QY 353 SRMLMWAKLYGKARKILHELFPSSFFFMEMEKYKLFAMEFVKYKQLOKEIYISDDHRSI 412
DB 367 ENILLQDQMWKAGRSILHMLMRTVPMIYDQKVRPAKAFMLHYNEIYEDFIKODHEMDV 426

Db 217 ESILQDEKTSRLSENKYGIDDDSCNNYSILVWMDKHSFKQFYEQITTA--ELPNNVFG 274
 QY 257 QLHTTAIDKEGRANVAGAYACQAEKEDIK--SHSENVSOHPHLHVEVLHSEIMAHQKF 313
 Db 275 KXVANIINDIGR-----ACIVTETNIKELLKIGOKLAQINLAVSRMSREDIFREES 325
 QY 314 ALRLGSMNKMIMSYS--SDFRQIFCOACLR----- 341
 Db 326 CAVILLEWADIAGSSICGRNYPSSVICKELVRPNWCGLNHSDLTFRLSLRLAUPAIVA 385
 QY 342 -BEPD-----SRNPCLISRLM----- 356
 Db 386 IDSPDIFLNEDHNSGSDTSSHMLTDESSHSRHWYPSNSLPDVLASYASRVAFDYFF 445
 QY 357 LWDKLYKGARKILHELIFSSPFMEMEYKLFAMFEPVKYKQLOKEYISDDHDSISITA 416
 Db 446 LYDLKLWKSRLYKLOELYGYFITOPGPKBEIMGARIAISYRRLAELFLLLDREPEHSVIF 505
 QY 417 LSVQMETVPTLARHLEEQNVISVITETLVL-----PEYLD-----RNNKFNFOGY 464
 Db 506 FSNQIPTADVAKGLVTEYDFUTINATLYTFTYKKLNTPNYVDQHAMIRTDAAFH-- 563
 QY 465 SODKLRGVAVICDLKYIILISKPTIWTIRLR--MOFLEGRSFLKILTCMQGMEIRROV 522
 Db 564 -----SRRYIHIFHHIQFMLSIPCV-AEIVREDLKFKQYADFENL---FQGMCPYTRAV 614
 QY 523 GQHIEVDPP-WRAAIAIQML-----KNILLMPQEW-----CACDEBELLVAYKECHKAV 571
 Db 615 SQHVEWENDSMYVNLVNSLQVAKLCHRGVGNVFMELNTKLANAINYLISLILYPKARNES 674
 QY 572 MRCSTSFISSTVQSCGHSLETKSYRVSDELVSITHLPLSRTLACGLHVRSLRGAVSRL 631
 Db 675 W-TNTESLTGTITVDER-GNS-KLIBYDIALQPVSFHHPLHMLVVL-----L 719
 QY 632 HEFVSPEDF-----QVEVL--VEYPLRCLVLAQVVAEMWRNGLSLISQVFFYQDVKCR 684
 Db 720 SPVVERDNYKLLWTQLDLAVTDHPLRVCAWLSQMBKAWIRNGTTLRQDAHHRNLSFH 779
 QY 685 BEMYDKDIIMQIGASIMDPNFKLLVLQORYELAEAF--NKYISTQODLIQYNTLIEE 742
 Db 780 EYTFDLVLLQLLTGYDPAIDLPISFRFOLEDQMGYRFFVPHKHYD-VSQVTIMMEE 838
 QY 743 MLOVLIYIGERVVPGVGNVKEEVTWRELIHLCLIEPMHSAIAKNLPENNETNGLN 802
 Db 839 FLILLISVNCNTAVLDHWDITRR--TEYGLAHLCPRLPSPYSITKRTCEHLLHKKQFES 896
 QY 803 VINKVATFKK-PGVSGHGVYKXDSLKDFNMFFYHYSKTQHSKAEHMOKKR-RKOENKD 860
 Db 897 TLKKVATFRNAEINDSGSFTLDEYFDYVDPNIIHYSRQREAEANILRRRYSKQHSKH 956
 QY 861 EALPPPPPPFCFPAFSK--VINLLNCDIMMYIL-RTVFERAIDTDSNLMTEGMLQWAFHI 917
 Db 957 --LESVVYBYEPHLSHNTIPIQLSDSFVGLWHTIIVVAYIYYPDQKLEGLVNTALHA 1014
 QY 918 LALGLLEEKQLOKAPAEVEVTFPYHKASRLGSSAMNIQMLLEKLGIPQLEQKD---- 973
 Db 1015 CLLVLMSK-----GSEPIFSKKICE--NRFPVVEGLQBYCNS 1050
 QY 974 ----MTTWILOM-----FDTVKRLRKSCLIV-----ATTSGSESINKDE 1009
 Db 1051 PDVTLSVLCQMKHNRNFVYVK--EKISLIMKILKSEVPLLYEPVVAETLSISSSKIYQ 1107
 QY 1010 ITHDKBAERK-RKAEARLHRQKIMASALQKNFIETHKLYMNTSEMPGKEDSIMEE 1068
 Db 1108 SLSDAEQOEQHLAKVRMAKERQARIMEQPRMQNKFLENHALPEASDCEN-----DEADEF 1163
 QY 1069 ESTPAVSDYSRIALGPKRGFSVTEKEVLTICILCOEBOEVKIENNAMVLSACTOKSALT- 1127
 Db 1164 SVTSSVS--TKJFLDP-----PIDTCLLQCEELKDRPYGTTLVF---VLRSSVRL 1209
 QY 1128 -----QHRGKPIELSG--EALDPL----- 1144
 Db 1210 FPADDANYSEVLDIPDLSLDEHIEQERFPFLAGKKKVKVLDSTEAYDYNYYEKKGNELHQ 1269

QY 1145 -----FMPDPLAYGTYTSCGHVMAHVCWKYFEAVQLSSQ---ORIHVDLFLDESGE 1194
 Db 1270 LKDSFNGPPDQDGLHATCGHFHMDCFKNHITATVTLATRANPYRHP--HNLSMKE 1327
 QY 1195 YLCPICKSLCMTVPIIPILOPKINSENADALAQLLTLARWI-----QTVL-- 1240
 Db 1328 FLCPLCKALCNTIFPILWRPKBEINFQAGVLT--APLKNMLVSKTSPFNKDLNQQLLDI 1385
 QY 1241 -----ARISGNIRHAKGENPIPIFFNQMGDSTLEFHSILSPGVESSIKYSNISKEMV- 1294
 Db 1386 ETPSPSEHTQSYNL-----NLDDV-LQHTLRDLSKDIYT-LNTGADNS--SDNVEENAD 1434
 QY 1295 ILFATTIYR-----IGLKVPDPDRPRVPMLTWSTCAF-----TIOAIENLLGDE 1339
 Db 1435 NLFQSSVLDHVFHFKSVNNEVPADER-----LAISDDIFELYRRLDDVDLANSLSYDD 1488
 QY 1340 GKPLFGALQNRQHNGLKALMQFAVAQ-----RITCPOVLI-----QKHLVRL 1382
 Db 1489 PIPVNGKL---HNVVK-LFSVSLCQVEASTRGHIKCSIPADIVWNLGKNQVFLRL 1543
 QY 1383 SVVLENIKSDETPCLLSIDLPHVLVCAVLAP-----PSLYWDDPV--- 1422
 Db 1544 S---ESIKTYTLLC--AHDQKRIGGSIQEFEFISFCQOKRIFGRLLPSL--DSPVTKSI 1596
 QY 1423 -----DLQPSVS-----SSYNHL-YLPHLITMAHMLQILTLVDTGLPL 1460
 Db 1597 TDDRVEPLLVKDTFREFABASVGLSCDESFPHYLTQLYTTADIYVNLWILLSQRNSL-L 1655
 QY 1461 AQVQSDSEEA-----HSASFFABISQYTSGSIGCDI-----PG--WYLWVS 1500
 Db 1656 KCMESVEFEAPDYEQKGFELVIQIWKSLRVGDAGLINFDCCTEDDLNPHLLFTLYKL 1715
 QY 1501 LKNGITPYLRCAALFFHYLLGVTTPPEELHTNSAESEYSA LCSYLSLPTNLFL--LFQBY 1557
 Db 1716 LERFSLILFRKCALLMYCRYGVSPETQPNLNPQNSLSRLQTKMHPGVIELSNHLCLTA 1775
 QY 1558 WDTVRPLQRWCADPALLNCLKOKNTVVRYPKR---NSLTELPPDYSCILNQASHFR 1613
 Db 1776 STESWLSLHKWC-----NFTETGPLCDFFRAYYPGIYELVSLPYELDKVPELLARAC 1829
 QY 1614 PRSADDERKHPVLCFCGAILCSQNICQEI VNGEVEGACIFALHCGAGVCFIKIRBC 1673
 Db 1830 SKCL-TEPMEPAICLFCGKLLCFQSHCCS--FNG--IGECNLHMQCCASDIDGIFLIVK 1884
 QY 1674 RVVLVEGKARGCAYAPYLDDEYGETDPLKRGKGNPLHLSERY-RKLHLVWQOHCIEETA 1732
 Db 1885 AILYLNPPV-GSFSVAPFLDAYGETDLGLRRGRSOLYSQRYDETVRTWMLNGSIPSYIA 1943
 QY 1733 RSQETNQMLPGFNWOLL 1749
 Db 1944 ROLDANPDGTGG--WETL 1958

RESULT 4

T37711
 probable n-end-recognizing protein - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C:Accession: T37711
 R:Murphy, L.; Harris, D.; Wood, V.; Barrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, June 1997
 A:Reference number: Z21738
 A:Accession: T37711
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-2052 <MUR>
 A:Cross-references: UNIPROT:O13731; EMBL:Z97208; PIDN:CAB10108.1; GSPDB:GN00066; SPDB:S1
 A:Experimental source: strain 972h-; cosmid c15A10
 C:Genetics:
 A:Gene: SPDB:SPAC15A10.11
 A:Map position: 1

Query Match			9.98; Score 914; DB 2; Length 2052;
Best Local Similarity			20.48; Pred. No. 1.2e-47;
Matches 447; Conservative 326; Mismatches 734; Indels 682; Gaps 89;			
QY	62	QBSVQMSIPTLEW-YLFGEDEP-----DICKELKHS-----GAFQLCGRVF 103	
DB	43	QEVFLSILNVDNWKYFLKKEGATSDPRLSRLQHSPECAQELQDKRSGSKVCGHVF 102	
QY	104	KSGETTYSRCDCAIDPTCVLCMDCFQDSVHNHRYKMHGTG--GGFCDCDGTBAWTKGPF 162	
DB	103	RAGEVIYRCNCGLDNNTCVLCAPCFRATNHEGTHVSISTSYSGICDCDGPANVNDLN 162	
QY	163	CVNHEPGRAGTIKENSRCPNBEVIVQARKIFPSVIKYVEMTIW-----207	
DB	163	CKIH-----NVPDDBEQKPEEVI-----PLEQHSIRTIHILLDFILVDFSCSPV 209	
QY	208	-----EESKELPPELQIREKNERY-----YCVLPNDEHSHYDHIYS 244	
DB	210	NLKAQSTVSGILADEBAS-----RUSSAKYGVADPCNVRVLMWDEVHTFDVAVGS 262	
QY	245	LOALDCELABAQLHTT-AIDKEGRRAVKAGA-----275	
DB	263	VLEALDSSNTAFGLVAQRVDSIGRPVATSASVHEAIRIANAIKENLAVNVRTARDPF 322	
QY	276	-----YAA-----COE-----281	
DB	323	REDICGILLEWFDLLESHVCYPADYLIIVCDEILKNWSPGLEKPAKPEVFNPLLEI 382	
QY	282	-----AKEDIKSHSENVS-----QHPL-----298	
DB	383	VNDDESDIYAAEELLDVIANIQDETGVTRIANLGGDEPDEADMTDPTIAGDFHPLDDD 442	
QY	299	-----HVEVLHSEIMAHQK-----ALRGSWMKNMSYS-----328	
DB	443	NDVNDLLDPETEREDIDDLTDEVMTEENEAADYGVNVRNTRQDDVDQDISMETESQNE 502	
QY	329	-----SDFRQIF-----335	
DB	503	TBESQNTENVYNPQTHPTVPIPTATQDVVTIRPEFNSQLNLRQIRNARRPRPAV 562	
QY	336	COACLREE-----PDS-----ENPCLISRL-----MLWDKLYKGARKILHELIFSSPF 379	
DB	563	CQVSLREDYKSPHPITPPSSYSVFESPSTLRDYFLLFDLKPWKLRLGLLSKLYVVPFN 622	
QY	380	MEMBYKCLFAMEVVKYKQLOKEIYISDDHDSRISITALSQVMTPTVTLARHLIEQNVIS 439	
DB	623	RNLLFKRLMGIRFVIHYRSLATAPLADREPDHSVMFLSVQPTTSLAEAVKVDYFLT 682	
QY	440	VITETLLEVL-----PEYLDNNKNFPQGYSD-----KLGRTVAVICDLK 480	
DB	683	NLMATTLISLTQSNRPSTL-----FSSDIETPTIQLNRQVLRKTRRTYNLPSDLG 732	
QY	481	YIL-----ISKPTIWTRELRMQPLEGRSPFLKILTCMOGMEETIRROVQGHIEVDPDWEAAI 536	
DB	733	YLLQHPQVKKLVDDTRYVHQYIDLLRVF-----QGVIPQORAILSHVQ-----WDFPH 781	
QY	537	AIQOMQLNILLMPQEWCADEEL-----LLVAYKECHKAVMRCSTSPSSKTVQSCGHS 592	
DB	782	G-----KNILFVMQVRVAMLSNVTSSCFTQAPYERLFYAI-KCIITSITHPKLDI-----AES 832	
QY	593	LETKS-----YRVSEDLVSIHLPLSRTLAGLHVRISRLGAVSRHLHFVS 636	
DB	833	LEPLSCIPSSSLTNFTQPLVPVPSRDPISFYHP-----LHWMLSNLFSYCRVDASSH 885	
QY	637	FEDQFQEVLEVPLRCLVLVAQVVAEMWRNGLSISQVFPYQDVCKREMYDKDILMLQ 696	
DB	886	WDKOTLALLDHLPRVCVLLAQIDCNLNRGSRISILLTDAFYRQLNIEVSYDVKDLAIQ 945	
QY	697	IGASLMDPNKFLILLQVRVELAB-AFNKTIISTKQDLIKQYNTLIBEMQLVLIYIGERY 755	
DB	946	TILMFVDPLNVLNAVQREFDMLNLTNYPNDYDTERIPAMLCMKLELLIALITER- 1004	
QY	756	VFGVGNVTKEEVTMBRIIHLCLIEBPHPSAIAKNLPENNETGLENVINKVATFKPK-G 814	
DB	1005	-BOILHVDIQQIIRTRLAQOLCRGPLAYSALLTSSNLVSESFPKIREEVTYSKAPDG 1063	
QY	815	VSGHGYELKDESLSKDFNMYFYHYSTQHSKASHMQGRKQEN--KDEALPPPPPEFC 872	
DB	1064	LHDPGVYSLKDEYDVLDPYFYHYNKEREESTILKRLAKKNVSAESIIEPKIRPL 1123	
QY	873	-----PAFSKVINLLNCDIMMYLRTVFERAIDTNSLWTEGMLQW-AFHILALG--LLE 924	
DB	1124	EKQGHDIPTFAAVNASTFSLIIP-----RAIE-----YALVQAESFGSDIGNITLG 1169	
QY	925	EKQOL-----QKAPEEBEVTFDFYHKA-----SRLGSSAMNIQMLLEKLGKIPOL 968	
DB	1170	DALQCLISMKIHEFSKSNDFCSRCAERYPTDSSIMRFBGGSAYCLAELCPAILKSPKY 1229	
QY	969	EQKQMITW-----LQWFD--TVKRLREKSLIVATTSGSEISKNDITTHDKAEARRKA 1023	
DB	1230	KOVHVKNVAVLAGLQKNDPAYSANMLEATHPELSTTSS-----TSDSNEIEKTQEK 1280	
QY	1024	EAARLHQ--KIMAQMSALQKNFIETHKLMYDNTSEMPGK---EDSIMBESTPAVSDYSR 1079	
DB	1281	KRLALEKQKKIMQOFRDQOASFLA-----QNTDFDIGEDQTEDEVTTBEPEBEVKYHEH 1334	
QY	1080	IALGPKRGPSTVEKEVLTICLQEEQEVK-----IENNAVLSACVQKSTALTQHRG 1131	
DB	1335	I-----RG-----NCLCOEBCNDQAPYGVIGIIOGSSLLRKTVDHSEIILDEIYS 1380	
QY	1132	KPIELS-----GEALDPLFMDPD-----LAY-----GTYTSCGCHVMHVAWCQKY 1171	
DB	1381	VPNLDRBESHSPFGKKYTVVFNRSKORLLSAYPGNNIRGVFVSGCHLMHLGCFKNY 1440	
QY	1172	FEAVQLSSOORIHVDLFDLESG--EY-----LCPLCKSLCNTVPIPIIPLQPKIN 1219	
DB	1441	YVARSMYRN-----DVTAGLSEYKYKYSTAKPFMCLCRSLSNVLLP-MEQIPKM-- 1489	
QY	1220	SENADALAQILLTARWIOVTLARISGYNIRHAKENPIPIFPNQMGSSTLEFHSI---- 1275	
DB	1490	CLNIDTLPFRSMNGWLEEI-GTMSSSFEYQ-----LVRSSLSDTKOTFRSCFLRP 1540	
QY	1276	-----LSFGVESSIKYSNS--IKEMVILFATTIVIRIGLKVPPDERDPRV-PM 1319	
DB	1541	WINSKIISAMLARLKIADGALIDQSNRVDSDLYDRYCETT-KLANKLVKGSTFTVNSPH 1599	
QY	1320	LTMSTCAPTIOAIENLLGDEGP-LFGALQNRQHNGKAL-MQFAVAQRITCPQVLIQKH 1377	
DB	1600	DLNLSLAYTVSSLE--VSQRCSPKQSGATRSVWFNSELGFLTSLFSLTSDT----- 1648	
QY	1378	LVRLLSVLVPNTKSEDTPCLLSIDLPHVLVGVAVLAPPSLY-----WDDPVDLQPSV 1429	
DB	1649	---VLKVCVDQI1KSDQOALLMESQKLAVCKI FYRHSOLKSMLRNGRMSDHDQIQPFL 1705	
QY	1430	SSSY-----NHLVFLPHLITMAHMLQILLTVDITGLPLAQVQED---SEE 1469	
DB	1706	SNTFDDFVKISSMLIFGQDNILTYVKLFYLSBECKTIISMIKKVVADSVVDPDLTINS 1765	
QY	1470	AHSASSFFABISQYTSIGSDIPGWYLMVSLKNGI-----TPYL 1509	
DB	1766	QOQSKQFY-----ILC--KNVLLWCGSSNNIEILDDESNNLLRLMSLVKEYSLPFL 1813	
QY	1510	RCALPFFHYLLGVTPP-BELHTNSAGEYSALCSYLSLPTNLFLLPQFW-----DT 1560	
DB	1814	RRVALVLYCMFDISLEFNEFBSNNEDDSELERLSKLIKVP--LQELYSQMSDENNQ 1868	
QY	1561	VRLQORWACDALLNCLKQKNT-----VVYPRKNSLIE-LPDDYSL 1604	
DB	1869	ILEIAGWCEHLA-----QNTWGDSTISLEYPGIYELVKLPHRLNIDSMSQSVCCM 1921	
QY	1605	LNQASHFRCPRSADDERKHPVLCPCGAILC---SQNICQBIIVAGEYVACIFIALHCG 1661	
DB	1922	CHKTPL-----PALCMLCGSVICFNARQNTVSSRLTGE---CNKHAATCT 1965	
QY	1662	AGVCIPLKIRECRVVLVEGKAR-GCAYPAPYLDYGETDGLKGNPLHLSRERY-KLH 1719	

Db 1966 GSVGIFFITKACGILLDLSINTGTTIMPTYPYLDIHGETDLQRRGCPQFLNQKRYDFVVR 2025

QY 1720 LWQOHCIIIEETARSOETNOMLFGFNQL 1748

Db 2026 EQWLRTVLQKMARHMDTMEQ---NWRM 2051

RESULT 5

S12332

ubiquitin-protein ligase (SC 6.3.2.19) - yeast (*Saccharomyces cerevisiae*) (strain S288C)

N:Alternate names: protein G7168; protein YGR184C

C:Species: *Saccharomyces cerevisiae*

C:Date: 16-Sep-1992 #sequence revision 16-Sep-1992 #text_change 09-Jul-2004

C:Accession: S12332; S64502; S64498

R:Bartel, B.; Wuening, I.; Varshavsky, A.

EMBO J. 9, 3179-3189, 1990

A:Title: The recognition component of the N-end rule pathway.

A:Reference number: S12332; MUID:91006011; PMID:2209542

A:Accession: S12332

A:Molecule type: DNA

A:Residues: 1-1950 <BAR>

A:Cross-references: UNIPROT:P19812; EMBL:X53747; NID:G4743; PIDN:CAA37779.1; PID:G4744

A:Experimental source: strain S288C

R:Arriyo, J.; Garcia-Gonzalez, M.; Garcia-Saez, M.I.; Sanchez-Perez, M.; Nombela, C.

submitted to the Protein Sequence Database, May 1996

A:Reference number: S64499

A:Accession: S64502

A:Molecule type: DNA

A:Residues: 1-1950 <ARR>

A:Cross-references: EMBL:272969; NID:G132325; PIDN:CAA97210.1; PID:G132326; MIPS:YGR18

A:Experimental source: strain S288C

R:Hebling, U.; Hofmann, B.; Delius, H.

submitted to the Protein Sequence Database, May 1996

A:Reference number: S64003

A:Accession: S64498

A:Molecule type: DNA

A:Residues: 1615-1950 <HEB>

A:Cross-references: EMBL:272969; MIPS:YGR184C

A:Experimental source: strain S288C

C:Genetics:

A:Gene: SGD:UBR1; PTR1

A:Cross-references: SGD:S0003416; MIPS:YGR184C

A:Map position: 7R

C:Superfamily: ubiquitin-protein ligase

C:Keywords: ligase; protein degradation

Query Match 7.5%; Score 689.5; DB 2; Length 1950;

Best Local Similarity 19.8%; Pred. No. 7.5e-34;

Matches 409; Conservative 323; Mismatches 732; Indels 601; Gaps 86;

QY 7 GGTREMEISABLPQTPQRLASWDDQVDFYAF-LHHLAQLVPEIYFAE-----MDPDL- 59

Db 33 GPTERADMSRALKEFIYR-----YLYFVINSGENLPTLFNAHPKQKLSNPFLT 81

QY 60 ---EKEESVQMSIFPLEWYLFGE-DPDCLEKLHSGAFQLCGRVFKSGEYTYSCRD 115

Db 82 VFPDSLEDAVDIDKITSQQTIPYKIDSRIGDVHKG--RNCGRKFKIGEPLYKCHEC 139

QY 116 AIDPTCVLCMDCPQDSVHKHRY--KMHSTGTGGFCDCGTEAWKTPFC-----VN 165

Db 140 GCDDTCVLCHCFNPKDHNHVVCTDICTFTSGICDCGDEEAWNSPLHCKAEQENDIS 199

QY 166 HEPGRAGTIKENSRCPLNEEV-IVQAKIIPSPVIVKVTWETIWEKEKLPPELQ----- 218

Db 200 EDPATNADIKEED--VWNSVNIALVELVAEVDYFID--VFQNIETPLTTIQDKITIK 255

QY 219 IREKN-----ERYVCVL 230

Db 256 LEWTOQGRMYERAQFLNDLYKENDYMFQDGTAKTSPNSPRASPLAKIDPENTVII 315

QY 231 FNDHEHSYDHVIYSLQALDCELAEAQLHTTAIDKEGRRAVKA-----GAYAAQCE- 281

Db 316 YNDEYHNYSQATTALRQGVF-DNVHIDLLTSRIDGGRAMLKCSQDLSVLGGFFAVQTN 374

QY 282 -AKEDIKSHSENVSOHPLHVEVL---HSEIMAHOKFALRLGSMWNKIMSYSDP----- 331

Db 375 GLSATLTSWSEYILHOETCKYIILWTHCLNIPNSSFOFTFRNMCKTL---CSEYLNATEC 432

QY 332 -----RQIFCOACLBEP-----DSN 348

Db 433 RDMTPVVEKYFSNKFDPKNDPYRIIDLSDAGNQIPLGHHKILPESSTHSLSPINDVET 492

QY 349 P-----CLISRLMLDAKLYKGARKILHELIPSSPFMEMEYKGLPAMEFYKYKQLOK 401

Db 493 PSTRYSNTRLOHILYFDNRYWKLKRDIONVILPTLASSNLYKDFICQOVVVEIFHNHTR 552

QY 402 B--YISDDHDSISITALS---VOMFTVPTLARHLIEBQNVLSVITETLLELPEYL--- 453

Db 553 SVAYM-----DREPQUTAJRECVVQLTCTPTNAKNIFENQSFIDIVW-SIIDIFKEFCKVE 607

QY 454 -----DRNNKFNF-QGYSQDKLGRVAVICDLKYLISK---PTIWTRELRMQPLEGPR 503

Db 608 GGVLIWQVRKSNLTKSYISFKGLYTV-----ETLLSKVHDPI---PLAPKEI---I 656

QY 504 SPLKILTCMQGHEETRRQVQHI-EVDPDEAAIAIQKMLLMFQEW-----ACDE 557

Db 657 SLLTLCKLFGCAWKIKRKEGEHVLHEDQNFISLEYTTSIYSIIQTAEKVSRKSDIDS 716

QY 558 ELLIWAYKECHKAVMRCSTSPISSSKTVVOSCGHSLKTSYRVSDLSVILPLSRTLAG 617

Db 717 KLFINA-----IRIISFIGNSLTYKLIYDHEVIFKFSVSHERVAFMPPLQTMLSF 768

QY 618 LHVRLSRIGAVSRLLH---EFVSFEDQVEVLVEYPLRCLVLVAQVVAEMRRNGLSLSIQ 674

Db 769 LIEKVSLEKDAVEALEDCSDFLKISDFS-----LRSVLCSDQDVGFWNGMSVLHQ 820

QY 675 VFYQDVKCREM---YDKDIIMLQGA--SLMDPNKFLLLVLRVELAEAFNKITISTKQ 730

Db 821 ASY---KNNPELGSYSRDIHLNQLAILWERDDIPRIIYNILDRWELLDWFTGEVDYQHT 877

QY 731 DLIKOYNTLIBEMQLVIVGCVVPGVNTKE---EVTREIHLICIPMHSATA 787

Db 878 VYEDKISFIIOQFIAPFIQILTERQYFTFSSLKRRMDQIKNSIINYLMFPLSYKLL 937

QY 788 KNLPEN-ENNETGLENVINKVATPKP-GVSGHGYVELKDESLKDFNMVYFYHYSKTQHSK 845

Db 938 RSPDYLTDTTFEDBALEEVSVFEPKGLADNGVFKLK-ASL-----YAKVDPLK 987

QY 846 AEHMQK-----KRRQENKQDALPPPPPEFCFAPSKVINLLNCDIM---MYLIT 893

Db 988 LLNLNENEFESSATIIKSHLAKDKDBIAKVVLIPQVS-----IKQLDKDALNLAGFTRNT 1041

QY 894 VFERAI-----DTSNLTWTEGMLQMAFHILALG-LLEEKOOLKA-----PEEVTFD 940

Db 1042 VFAKVYKLLQVCLMDEUSTFLNELHLVHGIFRDELLINGKDSIPEAVLSKPCINLLS 1101

QY 941 FYHKASRLGSSAM--NIQMLEKLKGIPOLEGOKDMITWILQMFDTVTKRLKSKCLIVAT 998

Db 1102 IANAKSDVFSEI VRKADYLLEKM-----IMKPNELFESLIA- 1139

QY 999 TSGSESIKNDIETHDEK-----AERKGAEAARLHQKIMAOHSAQKQNFIEH 1048

Db 1140 -----SFGQYVNDYDKKLQGVNLOETEKERKRLAKHQARLAKFNNQOTKPMKH 1194

QY 1049 KLMYNTSMPCKEBSIMBERSTPAVSDYSRIALGPKRGPVTEKEVLTCILCOEBOEVK 1108

Db 1195 ESEFD---EQDNDVHW-----GEKYSESDFTCALCQDSS--- 1227

QY 1109 IENNAWLSACVQKSTALTQHRGKFIELSGEALDP-----LFMDPDLA 1151

Db 1228 -STDPPVIPA-----YHDSPIFRPGNIFNPNEFMPMDGFYNDDEKQAYIDDDVL 1277

QY 1152 YG-TYTGSG-----HVMAHVCWKYFAVOLSSQRIHVDLPDESLEYLCLCK 1201

Db 1278 EALKENGSCGRKRVFVSCNHHHTHNCFKRYVQKGRPS-----SNAPICPLCQ 1324

QY 1202 SLCTWVPIPILOPQKINSADALAQLLTLARWIOVTLARISGYNIRHAKGENPIPIFF 1261
DB 1325 TFSNCTPL--CQTSKANTG-----LSLDMFLESEL----- 1355
QY 1262 NOQMGDSTLEFHSILSGVSESSIKYNSIKEMWILPATTIYRIGLKVPPDERPRVPMLT 1321
DB 1356 -----DTLS--RLFPKPTENTRTINSISLMI----- 1381
QY 1322 WSTCAFTIOAENLLGDGKPLGALONRQNGKALMQPAAQRITCPOVLI-----Q 1375
DB 1382 -SQCGQFDKAVRKANFPHKOVSLILSVHWANTISML---EIASRLEKPYISIFPRSRQ 1437
QY 1376 KHLVRLLSVLPNIKSEDTCLASIDLPHVLGVAVLAPPSLYWDDPVDLOPSSVSSYNH 1435
DB 1438 KY-----KTLKNI-----LVCIMLFTFVIGK-----PSMEF-EPYQOPQDVMNQ-NQ 1478
QY 1436 LY-----LFHLITWAHMLQILLTVDTGLPLAQVOEDSEASASFPABEISO----- 1482
DB 1479 LFQYIVRSALFSPVSLRQVTREALTTPSRQFLDRDFLOGLSDAQVTKLYAKASKIGDVLK 1538
QY 1483 -----YTSGISGCDIPGWLWVLSKNGITPYLRCAALF---PHYLL 1520
DB 1539 VSBQMLPALRTISDVMEGLDSESIYDLAYTLKSL-----LPTIRRCVLFKVLHEL 1594
QY 1521 GVTTPPELHTNSAGEVSALCSYLSPTNLFLLFOEWTVDVRLQRCWCAADPALLNCLQK 1580
DB 1595 KOSNETLVINGHEVEE-----LEFEDTAEFVNKALMITEKESLVDLAT 1641
QY 1581 KNTVVRYPKRN-----SLIELPDDYSCLLNQASHPCRSAD-----DERKHPV 1625
DB 1642 QBSIVSHPYLENIPYEGYGLIKLIDSKYLTNTVTSQKEIKLEERSQHMKNADNRDPK 1701
QY 1626 LCLFCGAIL-----CSQNICCOEIVNGBEVGACIFALHCGAGVCIFLK 1669
DB 1702 ICLTCGVKHLRADRHMTKLNKCFKPGAFMPNSSE-----VCLHLT 1747
QY 1670 IRECRVVLVEGKARGCAYPAPYIDYGETD-PGLKGNPNLHLSRERYKHLVWQHCII 1728
DB 1748 QPPSNIFPI-----GAPYLNSHGVEGRNMRGDLTTLNKRYEHLNRLWINNEIP 1797
QY 1729 EETARSQE-----TNQMLGPN 1745
DB 1798 GYISRVWGDGFRVTLISNGLFAPN 1822

RESULT 6
T30554
ubiquitin-protein ligase E3 - yeast (Kluyveromyces marxianus var. lactis)
C:Species: Kluyveromyces marxianus var. lactis, Candida sphaerica
C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
C:Accession: T30554
R:Waller, P.R.H.; Varshavsky, A.
submitted to the EMBL Data Library, April 1998
A:Description: Kluyveromyces lactis UBRI, the recognition component of the N-end rule pa
A:Reference number: Z20856
A:Accession: T30554
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-1941 <WAL>
A:Cross-references: UNIPROT:O60014; EMBL:AF061554; NID:g3114670; PID:g3114671; PIDN:AACI
C:Genetics:
A:Gene: UBRI1
C:Superfamily: ubiquitin-protein ligase

Query Match 7.0%; Score 647.5; DB 2; Length 1941;
Best Local Similarity 20.5%; Pred. No. 2.9e-31;
Matches 409; Conservative 319; Mismatches 689; Indels 581; Gaps 91;

QY 35 FYTAFHLHLAQVPEIYFAEMDPDLKQESVQMSIFTPLWYLFCEDPDICLEKLKHS 94
DB 51 YYYTMSDSGRLLPHFMFTATNERBFKNVDQAMEIKLSKPMYKIDENG--HSPFNHAG 108
QY 95 AFOLCGRVFKSGBTYTCRDCDAIDPTCVLCMDCFQDSVHKNHRY--KMTSTGGGFCDCG 152

DB 109 --RICAKFRVGEPIYRCKECSFDDTCVLCVNCFNPKDVGHHVYTSICTEFNNGICDCG 166
QY 153 DTEAMKTGPFVNHPEGRAGTIKENSK-----CPLNEEVIYQARKIPPSVIKY 200
DB 167 DKEAW-----NHELNCKGA-EDNGRLEDEPDHDKGKISKMLSVLIELPDPHIDVFNQ 218
QY 201 VVE--WTIWE-----EKEKLPELQIREK-----NER----- 225
DB 219 NIEPLTTIOKPLIAKLYFENPEREIVEQADMLRLAYRNQYMDDEBESSNKRHLTSLDPLS 278
QY 226 ---YCVLPNDHEHSYDVIIYSLQALCELAOLHTTAIDKEGRRAKAGA-VAACQ 280
DB 279 TLKDYAILYYDEFNYSQASAAIRQG-GPDNKHIDLLTAKIDSEGRSLRCSADIASLM 337
QY 281 EAKEDIKSH--GENVSQHPLHVEHSEMAHQFALRLGSMWNKIMSY--SSDFROIFCQ 337
DB 338 GRIPSVQSNGLCTTIQ---WTEYLHQEAC---KYSI---MWINDCLNIPNSTFQSLFRN 388
QY 338 A-----CLREEPD-----SENPCLI----- 352
DB 389 AIGKVLCSKYEPFYQSIDMTSVVRYDFSYSLSDDPYLVADHSVLGEGVKIPLGRHKSLD 448
QY 353 ---SRL-----MLWDKLYKGARKILHELIFSSFFMEMEYK 385
DB 449 PGDISAISPILNKVIAEDHHEYNLSRLQYVLFENRYWKGKARKIVQDLIPTLASSAVQK 508
QY 386 KLPAMEFVKYLOKOKEYISDDHDSISITALS---VOMFTVPTLARHLEE-----Q 435
DB 509 PHMTDOLVEIPHMTRS--GTFMDREPQUTLSRESVQULFTCTPTAYSIFHSGHFNLIW 566
QY 436 NVISVITE--TLLE--VLPEYLDNRNNKFNQGYSDKLGRRVAVICDLKYI-----LISKP 487
DB 567 SVIDVPDPSFTDEGTLVWQVRQSRNP--SKYSISIFKQGLYAVETLSKITDPNLLKLP 624
QY 488 TIWTERLRMQFLEGFRSFLKILT---CMQGMEEI--RRQVQGH-I-EVDPDWEAAIAIQMQL 542
DB 625 -----GEFIMIVTCLKLFNGAWKIKKREGHEVLRDQHFIPYLEYTTVS 669
QY 543 KNILLMF-----QEWACADEILLVAYKECHKAVMRCSTSFSSSKVTQVQSGH----- 591
DB 670 YSIQTDFKVLQOSKDHIDQRLIGA-----INLDSFL-----GHRNLSYK 711
QY 592 ---SLETQSYRVSEDLVSIHLPLSRLTAGL--HVRLS-RLGAVSRILHEFVSFDFQVEVL 645
DB 712 LYKDPELIKFOJISKEQVSPMPVHTLFCFLVQHVPLQVSIQVLSQSKDYLIVISDF----- 766
QY 646 VEYPLRCLVLAQVVAEMRRNGLSISQVFTYQDVKCREEM--YDKDIIMLQIGASLMD 703
DB 767 ---ALRSVVLCSQIDIGFVWRNGMSVLHQSAYY---KNPPEMSSYSRDIQLANQL-AFLIE 819
QY 704 PNKELLV--LQRYELAEAFNKTIQKQDQDLIKQYNTLIEEMQLVLIYVGERVYVPGVG 760
DB 820 KNDFORVIYNMLDRWELLDWFGVSPSTETVYDDKISSIIQQFVAPLYQLLVER----- 873
QY 761 NVTKEBVTMR-----IIHLACIEPMHSAIAKNLPEN-ENNETGLENVINKVATP 810
DB 874 DPKKFDLLEETQLYKNKNAIYKLYAEPLSYTDLLNDIPDYLTVESVSQFDVLEEVSY 933
QY 811 KKPQVSGHGVYELKDESLKD-----FNM--YFYHYSTQHKSAKHEMKQKRRKQENKDBA 862
DB 934 IEPKGSKIWCQTKEGTLQENRFLNNGNDFEHSATIVKSHLADSKEKRAKIIKVPQ- 992
QY 863 LPPPPPPPCPAFPKVINLLNCDIMMYILRTVFERAIDTSDSNLWTEGMLQMAFHIALGL 922
DB 993 -----LLELDLDPGAR---ELGSPFRTNLFA---KLIFKLKLAV 1027
QY 923 LEEKQLOKAPREEVTFDFYHKASRLGSSAMNIQMLEKLKGPIDEGOKDMITLQMF 982
DB 1028 -----SDSSFSFTYELLHLIHAIFRDEENVNGKDSLPEAYISKPICDLLLSIV 1075
QY 983 DTWKLREKSCLIIVATTS-----GSESTKNDSITH-----DKE 1015
DB 1028 -----SDSSFSFTYELLHLIHAIFRDEENVNGKDSLPEAYISKPICDLLLSIV 1075
QY 983 DTWKLREKSCLIIVATTS-----GSESTKNDSITH-----DKE 1015

Db 852 ERRPGLSTA---ESLREBILPKLATGDPHTHSQVLSLPRDLSEKDELQVLDVSVYCN 908
Qy 813 PGVSGHGVYELKDESLKDFMYFYHYSTKQHSKAEHMQKRRQKQENKDEALPPPPPEFC 872
Db 909 PSGMNOQKYSQSSCKWELDY-----HPRMQSRDLQSAEERF-----SRYC 950
Qy 873 -----PAPSKVINLNCDDIMMYLIRTVFPERAIDTDSNLWT---EGHLMQ 913
Db 951 GVSALTTLQPRRMVYPPPLKGLARIGTKATFOIISALYALQSGTSVKSRAPDGLVIT 1010
Qy 914 AFHILALGLLEEQQLQKAPBE---EVTDFYHKAS-----RLGSSAMNIMLLEKLG 964
Db 1011 ALQLLSLSDICTQORQSNODCCLENSIPLELAGLEIIGIAQGTKEKESLLSLVSLMK 1070
Qy 965 IPOLEGQK-----DMITWI---LQMPDTV---KRLREKSLVATTSGSESINDE 1009
Db 1071 TRMGDRGHPPEPGSCNMISSWIGLLKCFSAIDSVCMNLIQSLAPEVVGSGDFKVMGS 1130
Qy 1010 ITHDKKAEKRKAARLHRQKIMQMSALQKNFIETHKLMYDNTSEMPGKE-----DSI 1065
Db 1131 TSDEKRAKAKERQAA-----IMAKMKAEQSKFLSTLSSMDD--DPRSEFETSDSV 1181
Qy 1066 MEEESTPAVSDYSRIALGP-KRGP-----SVTEKEVLTC 1098
Db 1182 MEHDSIAVREVCSLCHDPDSDKDPVSLFPLQVSGMTDAIICDC 1225
RESULT 8
S64851
probable membrane protein YLR024c - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein L1730
C;Species: Saccharomyces cerevisiae
C;Date: 01-Aug-1995 #sequence_revision 24-May-1996 #text_change 09-Jul-2004
C;Accession: S64851
R;Obermaier, B.; Piravandi, B.; Rinke, M.
submitted to the Protein Sequence Database, May 1996
A;Reference number: S64845
A;Accession: S64851
A;Molecule type: DNA
A;Residues: 1-1872 <OBE>
A;Cross-references: UNIPROT:Q07963; EMBL:Z73196; NID:gl360331; PID:e245500; PID:gl360332
A;Experimental source: strain S288C
C;Genetics:
A;Gene: SGD:UBR2; MIPS:YLR024C
A;Cross-references: SGD:S0004014
A;Map position: 12R
C;Keywords: transmembrane protein
P;189-205/Domain: transmembrane #status predicted <TM1>
F;804-820/Domain: transmembrane #status predicted <TM2>
F;889-905/Domain: transmembrane #status predicted <TM3>
Query Match 6.0%; Score 549.5; DB 2; Length 1872;
Best Local Similarity 19.8%; Pred. No. 3e-25;
Matches 388; Conservative 321; Mismatches 729; Indels 521; Gaps 94;
Qy 99 CGRVFKSGTTCSCDAIDPTCVLCMDCFQDSVHKHRY--KMHSTGCGFCDCGDTSEA 156
Db 98 CTRLCPPEITYYCFTSTNPLVEICELCFDKEKHVHVSVAKVVRPEGRICHGD--- 154
Qy 157 WKTGPPCVNHPGRAGTIK-----ENSRCPINEEVIQVARKIPPSVIKYVEMTI 206
Db 155 ----PPAFN-PPAFN-DPSAPKCKNELNIPISNDNSNVTDENVI-----SLNLYVLDLIDTV 205
Qy 207 -WEEKEL-----PPELQIREK-NERYTC-VLPNDHEHSHYDIYS----- 244
Db 206 SYKEAEASHSERKASSLMHPQNSITDDIMEKHECEPLVNDENFVFDNNNSNTRKEAH 265
Qy 245 LQALDCELAELAQHTAIDKEGRRAVKAGAYAAQEAQEKDISH-----SENVSQ-H 296
Db 266 MEWAQIEECNCHVMDLASTITRLNTPVEVAISITALESDHVDVTVLQSEPFETD 325
Qy 297 PLHVEVLHSEIMAH-----QKPAIRLGSMWNI 324

Db 326 QIAKEPQKENVVHVVRKADDIIFRKKTUDDLTDWLVSCLCFKAATSLQNKYALRI-SMLDVW 384
Qy 325 MSYSSDFRQIFQACILREBPDSNP-----CLISR----- 354
Db 385 YSHFSK-----MRVSPTNTNPDPSKINLLGGFLISNEDSDSWPKPWSLENIEDER 435
Qy 355 ---LMLWDAKL-----YKGAR-----KILHELFPSSP 378
Db 436 ISKILTYNERLITRAHSPNTVSHFYGSRFQYIIINSINILSKSKFKML-KIMASLF 494
Qy 379 FMEYKYLKAFMEFVKYKQLOKEYISDDHDSISITALS--OMFTVPTLARHLIE-- 433
Db 495 SIRDESRRFLAQAQYIDVLSVLYDAVSD-AKECQVTLMSILQYTFQDPSIANWTISSG 553
Qy 434 --EONVISVIT-----ETLLEVLPEYLDNRNKNFQGY--QDKLGRVYAVIC--DLKY 481
Db 554 FIERTIRPAPTLMAPNPEDLMSVLPISL-----YNGFKLPTETIRNRRTIICPKDLCT 606
Qy 482 IL-----ISKPTTWTELRMQFLEGFRSLKILTCMQGMEIRRQVQCHIEVDP-DWEAAI 536
Db 607 IMSANTVPBELLSNEAIFNAIIESFBSFNVL-----PLKRETKHEVEVENFDFS--- 656
Qy 537 AIQMOLKNILLMFOEW-----CACDEBELL-----VAYKECHKAVM--RC 574
Db 657 AFYFFESSILIMTGTTRSISLVKDAAFKQIVLKLDDVAQTRFESLTSNRKATSPDNA 716
Qy 575 STSFSSSKTVVQS-----CHSLETYSRVSEDLVSIHLPLSRLTLAGHLVRLSLG--- 626
Db 717 STNENDSNKATLSTVRETTICNVYVAETINQVQNTQYFPNPM--YLFKFIQMSQGRYE 775
Qy 627 -----AVSLRHFVSPEDPQVVLV-EYPLRCLVLVAQVVAEMRRNGSLISQVPPYQD 680
Db 776 PIPASLTNYINLFEVFDQRKALYISESALSTLVLTGINVGVFWVRNGTPIITHQARMYTK 835
Qy 681 VKCREMYDKDIIMLQIGASLMDPNKFLLLVLQRYELAEAPNKTISTKODDLIKQNTYLI 740
Db 836 YSMREFTYISDFNVQFQSMAMCNPDLMVYLSRMGLKHWAN-GVPMYDPTDETTVAVV 894
Qy 741 BEMLOVLIYIGE-RYVPCGVGNVTKE--EVTMR-BIHLILCIEPMHPSIAKLNENNN 796
Db 895 NECILLLQTLTEVRS--VMKSSKEGFERTPKSHIHALCFDTCYAIQVNCIPEHITK 952
Qy 797 ETGLENVINKVATFKKP-GVSGHGVYELKDESLKDFMYFYHYSTKQHSKAEH-----M 849
Db 953 HPSFDIYLEKYANYTSPVSLTDNGIFVLKEKYKDEIDPYIIGLSSRRYDVEKIRLNMA 1012
Qy 850 QKRRKQENKDEALPPPPPEPCPA-----FSKVINLNCDDIMMYLIRTVFERA 898
Db 1013 NLKKMKYED-----TFVPAKKVKDLKNTLFGSLYSISSVNTFGLFKNTLDHI 1061
Qy 899 IDTDSNLTMEGLQMAFHILALGLLEEQQLQKAPBE---VTFDPYHKASRLGSSAMNI 955
Db 1062 IKYD-----YDNLPRVHLIHLCVVNNLNEFNGILWHEVAIVDTFCHVHS-IGS--ILY 1114
Qy 956 QMLBKLKIGIPOLEGO-KDMITWIIQMPDTVVKRLREKSLVATTSGSSI-----KNDEI 1010
Db 1115 YCLLK--DNFSSHGKIREIFRYLM---ETAPHVNVNSYLREQTTSYTPGILWPTKEDKS 1169
Qy 1011 THDKKAEKRKAARLHRQKIMQMSALQKNFIETHKLMYDNTSEMPGKSDSIMEEES 1070
Db 1170 HKDK---BPERKKHARLARKKLMKLAQQAQQMKFMENNSV---DTSDI-----S 1212
Qy 1071 TPAVSDYSRIALGPKRGPSTVEKEVLT-----CILCO--BEQEVKI-----ENNA--- 1113
Db 1213 TPRTTSPS---LSPTRNAENSSNTINSCDDDCVCKMPKDDDDVFVTFVSQYERNICDHG 1269
Qy 1114 -----MVLACVQKSTALTQHRKPIELSGEALDPLFMDPLAYGTTTSGCG 1160
Db 1270 IDFTNPTDVRINSLFSKGQTKDSAIQEN---PODDGDTRLKFTCEPVLV-----ACG 1320
Qy 1161 HVMHACVQKQFEAVQ-LSSQRIHVDFDLSESGYLCPCLKSLCNTVIP----- 1209
Db 1321 HGSHTKCLSGHMKSIIRGIONOTTKNIP--SYSGGLIYCPVCNLSNLSPLKNTDIDKRTS 1379

QY 948 -----LGSAMNIOMLEK-----LKGIPOLEGOKMIT-----WILQMP 982
Db 1279 LIKREMETETLVCFAFIRLLTKETRKNGALVWVFKGILSGEYDKERVTGGRMYLARFV 1338
QY 983 DTVKRLREKSLIVATTSGESIKNDIBETHDKABE-----RKRKAABALHRQKIMAO 1036
Db 1339 TILTSLFVARQIIEGKXBEELRISKHSRQEHOKAPMPDPVKAABEAKRMEALM-- 1396
QY 1037 MSALQKNFIETHKLMYDN--TSEMPGKEDSIMBESSTPAVSRIALGPKRGPSTVEK- 1093
Db 1397 QNSAKSKAQWELKMTKEGMDTDAENVKVPDPSQQRKV-----YECPIGGEQNAPTVNP 1451
QY 1094 -----EVLITCILCOEEOVKIENNAWV-----SACVOKSTALTQHRGKPIELSGEA 1140
Db 1452 FGLAKLSTNFICEEQDASINTIDILLKDFEYHVSANLQSETRRPFPSKRRQATFEN 1511
QY 1141 LDPLFMDPLDLAGTYTSCGCHVMHVCWQY-----PEAVOLSSQORIHVDLPD--- 1189
Db 1512 QDIVKVNPLV-GTDLTKTCHAAHACFNAYRASLVGSFOEKKLENEK-----KYDGHOR 1566
QY 1190 -LESGEYLCPLCKSLCMTVIPI-----IPLOPKINSENADALAOQLTLARWIQTVLAR 1242
Db 1567 STDREVGCDMCRVTNVNIVPMSFDKPYTPIKTPASPMYSDV-----N-----R 1611
QY 1243 ISGYNIRHAKENPIPIPFQMGMDSTLEPHSILSPGVESIKYN--SIKEMVILFATT 1300
Db 1612 VMDVVLKAKG---PVQED-----ERNVKYSTNYSTREGGLFELY 1650
QY 1301 IYRI-----GLKVPDERDPRVPM-----L 1320
Db 1651 IGRHSADLAKRSQORCTVSTLMVSLAVVMYERSIMRKMGPERRKNQRNMTHELM 1710
QY 1321 TWSTCA-----FTIQAIENLGDGKPLFGALQNRHNGLKALMQFAVAORITCPQVL 1373
Db 1711 TASVATSKVDVFDVALSALTNLFAKTENSFPSPPEQSTSEPAQCTEAPN---PEAV 1767
QY 1374 IQKHLVRLLSVLPNTKSEDTPTCLLSIDLPHV-----LVGAVLAPPSLYWDDPVDIQ 1425
Db 1768 VGLSSDEMAAMITKPKRKOTPDCLSKPLPALDPKATLVRLAVLDNQSLTKDIQREIA 1827
QY 1426 PSSVSSSYNHLIYLPHLITMAHMLQIILT-----VDGLPLAQ-----VQDSERAHSA 1473
Db 1828 QNMIFASLGVSSYTLTCL-----ILRTGEKISALNKGEPKIQGLSDHLQSEAEACQA 1882
QY 1474 ----SFFAEISQYTSIGCDIPGWVYVSLKNGITPYLRCA-ALPHVLLGVTP-PEE 1527
Db 1883 LTYNTYFPHLAORLB-SPDVEPTDEYICKTWNCLIEPLRFYSYELLFHCNIGFDVNNQ 1941
QY 1528 LHTNSAEGE-----YSALCSYLSLPTNLFLLFOEYWD-----TVRPLLQRCADPALL 1575
Db 1942 IHSRNDLDNLVRLVGINAQNVPCG-----ANYWTRAFRQTLTMAKLR-IPQPCIV 1995
QY 1576 NCLKQKNTVVRPRKNSLIEUPDDYSCLLNQASHFRCPRSADDERKHVPLCLFGAILC 1635
Db 1996 BPLAWK-----PRR---ILKPTNPDELFGRYFHREC-NKCSSVPSPVMCLFCGEIIC 2045
QY 1636 SONICCO-----EIVNGEEVGACIFHALHCGAGVICILKIRECRVVLVEGARGCA 1686
Db 2046 -LNDCCRMHQHSGSDRVISMSEIEA---HABDCSSSGLFSLVSTSSVMVVSRGK-QAAI 2100
QY 1687 YPAPYLDYGETDPGLKRNPLHLRSRVRKHLVWQOHCITIEETARSOE-----TNOMLF 1742
Db 2101 WGTIYLDAAKEDRNLRKPKPLFLCESRLKWLEYDAE-----QEWQRPQWFMNTNSQAF 2156
QY 1743 GFN 1745
Db 2157 TSN 2159

RESULT 10
T48252
eceriferum3 (CER3) - Arabidopsis thaliana

N;Alternate names: protein T1E22.70
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C;Accession: T48252
R;Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancroft
submitted to the Protein Sequence Database, March 2000
A;Reference number: Z24489
A;Accession: T48252
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-795 <REV>
A;Cross-references: UNIPROT:Q9LZ94; EMBL:AL162874
A;Experimental source: cultivar Columbia; BAC clone T1E22
C;Genetics:
A;Map position: 5
A;Introns: 203/2; 220/3; 428/1; 488/3; 666/2; 701/2; 723/3; 749/3; 772/3
A;Note: T1E22.70

Query Match 3.7%; Score 343; DB 2; Length 795;
Best Local Similarity 23.0%; Pred. No. 4.6e-13;
Matches 190; Conservative 123; Mismatches 274; Indels 240; Gaps 43;
QY 1033 INAAQS-----ALQKNFIETHKLMY-----DNTSEMPGKEDSIMBESSTPAVDSRTA 1081
Db 56 LMLQLSDDTISANMIESIKARLIGNGQTEKRSSDGRGKDESNMB---SLEIAMYQTVR 112
QY 1082 LGPKRGPSTVEKVELTCTILCOEEOVKIENNAWVLSACVOKSTALTQHRGKPIELSG--- 1138
Db 113 -----NKIENWQSLTRVDHQPEAEN-----CSEKNSV-----GGSTLOGRPP 153
QY 1139 -----BALDPLFMDPLAYGTYTSCGCHVMHVCWQYFEAVQLSSQORI 1183
Db 154 DIRSRQTSRRPDAGSDGPHPIDCD-----GVYLLSCGHAVHQSLERYLSKLSKERSGRT 208
QY 1184 HYD-----LPDESGEYLCPLCKSLCMTVIPIIP-----LOPO-KINSENADALAOQLT 1231
Db 209 VFEGAHIVDLKKBKPLFCVCRRLANSVLPECPGDLCSVKLQDSPRTKLRRKDLQPSL- 267
QY 1232 LARWITQTVLARISGVNIRHA-----KGENPIPIFFNQMGMDSTLEPHSILSPGVES 1283
Db 268 -----WLSALCL-----LRSAAEVIIDGDRGKTVP-----QGDGPRR----- 301
QY 1284 IKYSNSIKEMVILF-----ATTIYRIGLKVPPDRDPVPMLTWSTCAFTIQAENLIG 1337
Db 302 -KOLASVSKMLWFPYPPKPEDKTKLKLWL--PPQS-----IVMMDTLKYSILSME--IG 350
QY 1338 DEG-----KPLFGALQNRHNGLKALMQFAVAOR-----ITCPQVLI-QKHLVRL 1381
Db 351 TRFAKNSMLPVYCIDSLYEELKTSKGTILSVLLRVVQSRTKNTIHRQRFVGMKHLAES 410
QY 1382 LSVVLPNTIKSED-----TPCLLSIDL-----PHVTVGAV--LAPPSLYW 1418
Db 411 ICYGVSSSSSSSIFGSEGTGSLKNIDLLMNRASDPVLAHDPFSSAMWALFCLPFPFLTC 470
QY 1419 DDVDVLDQPSVSSSYNHLIYLPHLITMAHMLQIILTVDGLPLAQVQEOSEEAHSASSPPA 1478
Db 471 EESL-----LSLVHIFHSVLSVQTVIAYACRP-----SELSELNFENLNL 512
QY 1479 EISQVTSGS-----IGCDIPGWVYVSLKNGITPYLRCAALFPHYLLGVTPPE 1526
Db 513 DISNALRESGGWEPFRSNMNLSCDIKD-----TIRKYSLPFLRCALLKWL--KSTPR 565
QY 1527 ELHTNSAEGEYALCSYLSLPT--NLFLLP--QBYWDTVRPQLQWCAADP--ALLNC-LK 1579
Db 566 KLH-----EESDMFDLPDSDPTTNDMDFIYSPQSELNHHVQELKMFNIPPTIDILNDELL 619
QY 1580 QKNTVW-----RYPRKNS-----LIELPDDYSCLLNQASHFRCPRSADDE 1620
Db 620 RSSTQIWLQHFQREYRNVNRKRSICITPVFPQLMKLPNLVQDILLQRCIKRCV-NCTKV 678
QY 1621 RKHPVLCFCGAILCSQ--NITCCQIYNGEEVGACIFHALHCGAGVICIFLKIRECRVVLV 1678
Db 679 IEEPVLCLCGS-LCSPINWSPCCRE-----SGCPNHATCGAGTVGLLIRR--TTILL 729


```
QY 1055 TSEMPGKEDSIMEESTPAVDYSRIALGPKR-----GPSVTEKEVLTLCILCOBEQEVKI 1109
Db 2127 TKSMSSLDQ-----DRDRVIDEAKWEQRFDAIQTKB-----BEVRLKE 2166
QY 1110 EN 1111
Db 2167 EN 2168

RESULT 13
A56539
giantin - human
N:Alternate names: macrogolin
C:Species: Homo sapiens (man)
C>Date: 19-Oct-1995 #sequence_revision 26-Jan-1996 #text_change 09-Jul-2004
C:Accession: A56539; S37536
R:Seelig, H.P.; Schranz, P.; Schroeter, H.; Wiemann, C.; Griffiths, G.; Renz, M.
Mol. Cell. Biol. 14, 2564-2576, 1994
A:Title: Molecular genetic analyses of a 376-kilodalton Golgi complex membrane protein
A:Reference number: A56539; MUID:94187728; PMID:7511208
A:Accession: A56539
A:Molecule type: mRNA
A:Residues: 1-3259 <SES>
A:Cross-references: UNIPROT:Q14789; EMBL:X75304; NID:g405714; PID:CAAS3052.1; PID:g4057
C:Genetics:
A:Gene: GDB:GOLGB1; GCP; GCP371
A:Cross-references: GDB:454958
A:Map position: 3q13.31-3q13.31
C:Superfamily: giantin
C:Keywords: coiled coil; Golgi apparatus; transmembrane protein
F:3238-3254/Domain: transmembrane #status predicted <TMN>

Query Match 2.1%; Score 190.5; DB 1; Length 3259;
Best Local Similarity 18.0%; Pred. No. 0.0096;
Matches 221; Conservative 207; Mismatches 424; Indels 377; Gaps 49;

QY 164 VNHEPCRGATIKENSRCPLNEEVIQVQARKIPFSPVIVVEMTWEEKEKLPPELO--IRB 221
Db 12 VLHLSGDDTDQNMRAPIIDPELQHSDEMFNNTTQEDVQERLAYAE-QLVVELKDIIRQ 70
QY 222 KNERYYCVLFNDBHSHYDVIVYSLOALCELAEAQLHT-----TADK--EGRAVKAGA 275
Db 71 KOVOL-----QKDEALQBERKAANKIKKLHAKAKLTSINKYIEKWAQGGTV 121
QY 276 YAAQCBAKEDIKSHSENVSQPHLVHVLHSEIMAHQKPAFLRGLSWMNKIMSYSSDPQRI 335
Db 122 LPTEPQSEQLSKHDKSSTEEEMIEKIKHLQKEELISTL----- 163
QY 336 QOACLBEPDSENPCLISRLMLWDKLYGARKILHELIFSSFFMEYKYLPAWEPVKY 395
Db 164 -QAQL-TOAQAEQPAQSS-----TEME-----EFVNM 188
QY 396 KYOLOKEYISDDHRSISITALSVQMFTPTLARHLIEQN-----VLSVIT 442
Db 189 KOOLQKB-----EEFISTLQQLSQTOBAQAQVVRKEDARFETQVRLHDELLQVLT 242
QY 443 ETLLVPLPYLRNNKFNFGQSQDLGRVYAVICDKYLKISL-----KPTIWTIERLM-- 496
Db 243 QADVETEMQQLRVLRQLKEEESLVGR--AQVDLLQQLTAAEQRNQLLSQQLQOME 300
QY 497 -----QPLEGRSPFLKILTCQGMETRRQVGHIEVDPDWEAAIAIQMLKNILAMP 549
Db 301 AEHTLRLNTVTEREESKLLKLEVEVAERKLSFH-----NLQEMHLLLEQF 349
QY 550 QEWACADDEL--LLVAYKECHAVMRCSTFI--SSSKT--VVQSCGHSLETYSYRVSDEL 604
Db 350 EQAGQAQAELEGRYSALQKHAEMEEKTSHTLSLQTKQELQSCADKALQKNSKLLQDK 409
QY 605 VSHPLPLSLTLAGHLVRLSLGA-----VSRL-----HEF-----VSPEDFQVEVLVEPLRC 652
Db 410 NEQAVQSACTIQLELDQLOQKSKESISQFLNRLPLQOHEFASQTSFPDVNTEGTAVTEEN 469
QY 653 LVLVAQVVAEMWRRNGLSLISQVFFYYQDVKCREEMVDKDIIMLQI-----GAS 700
```

```
Db 470 IASLQRRVVELENBKGALLSSI-ELEELKAENEKLSQITLLEAQNRTEADREYSEIS 528
QY 701 LMD-PNK-----FLILVLRQYELAEA 720
Db 529 IVDIANKRSSSAEESQDVLENTFSPQKHKELSVLLLEMEKAEQAEIAFLKLQLOQKAEAA 588
QY 721 FNTKISTKQODLIKQY--NTLIEMLQVLIYIVGRY----- 755
Db 589 DHEVL---DQKEMKQMEGHGIAPIKMKVPLEDTGQDFPLMPNBESSLPAVEKQAEQASTHQ 645
QY 756 -----VPGVGVTKKEVVTMREIIRIHLICIEPMHSA-- 785
Db 646 SRTSEISLNDAGVELKSTKQDGDKSLSAVPOIGQCHODELB-RLAKSQILELNLPHKQAQ 704
QY 786 --IAKNLPENNETGLENVINKVAFPKPGVSGHGV-----ELKDESL---KDFNNYF-- 835
Db 705 EYIEKNLDEKAEISNLNLQIEB---FKGNADNNNSAFTALSEERDQLLSQVKELSMVTE 761
QY 836 --YHYSKTOHSAEAMQKGRKQENK---DEALPPPPPPFPCFAPSKVINLLNCDIMMYI 890
Db 762 LRAQVKQLEMLNLAEAERQRLDYESTQAHNLLTE-----QIHSLSTEAKSKDVKIEV 814
QY 891 LRTVTFE-----RAIDTDSNLTWTEGMLQWAFH 916
Db 815 LQNELDDVQLQFSRQSTLIRLSQLOQNKSEVLEGAERVRHISSEKVELSQALSQKLE 874
QY 917 ILALG--LLEEK---QOLQKAPEE-----EYTFQPYHKASRLG---SSAMNIOMLLEK 961
Db 875 IYKMQLLLEKRDVETLQOTIEEKDQOVTEISFSMTEKMQVQNKREKFLGLVETIKLTKQ 934
QY 962 LKGI-----POLEGQKDMITWILQMEDTVKRLREKSLIVATTSSESINKNDEITHD 1013
Db 935 LNLLSRAEAKKEQVEEDNEVSSGLKNQYDEM-----SPAGQISKEELQHE 980
QY 1014 -----KKAERKKAABAARLHROKMAQMSALQKQPIETHKLMYONTSEMPGKEDSIMEE 1068
Db 981 FOLLKKEQQRKLOQALINNKELLQVRSLRE---ELANLKDESKKGIPLSETERGEV 1037
QY 1069 ESTPAVDYSRIALGPKGPSVTEKEVLTLCILCOB-----EOEYKIENNAWVLS- 1117
Db 1038 EEDKENKEYSE-----KCVTSK-----CQIEIYIKQTISEKEVELQHIRKDLSE 1082
QY 1118 --ACVQKSTALTOHRKGFIELSGEALDPL 1144
Db 1083 KLAEEQFOALVKQMNTLQDKTNQIDLL 1111

RESULT 14
A57013
early endosome antigen 1 - human
N:Alternate names: endosome-associated protein
C:Species: Homo sapiens (man)
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: A57013; S44243
R:Mu, F.T.; Callaghan, J.M.; Steele-Mortimer, O.; Stenmark, H.; Parton, R.G.; Campbell, J. Biol. Chem. 270, 13503-13511, 1995
A:Title: EA1, an early endosome-associated protein. EA1 is a conserved alpha-helical
A:Reference number: A57013; MUID:95286647; PMID:7768953
A:Accession: A57013
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1410 <RES>
A:Cross-references: UNIPROT:Q15075; GB:L40157; NID:g1016367; PIDN:AAA79121.1; PID:g10163
R:Seelig, H.P.
submitted to the EMBL Data Library, April 1994
A:Reference number: S44243
A:Accession: S44243
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-254, 'C', 256-257, 'LO', 260-276, 'A', 278-283, 'A', 285-519, 'D', 521-574, 'EQ', 577-
A:Cross-references: EMBL:X78998; NID:g475933; PIDN:CAA55632.1; PID:g475934
C:Genetics:
```



```

699 QY -----ASLMDNKFLLVLQRYELA-----EAFNKTI$TKDODLKQYN-----737
    :      |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
648 DB ABAEKREALWEKNTLELAQLQAEBAGAELOADLEDIOBEKEBIOKCLSESHOQEAAAT 707
    :      |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
738 QY TLIEEMLO-----VLIIVGRRYPGVGNVTKEBTWREIITHLCIE-----PMP 782
    :      |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
708 DB TQLEQLHQAKRQEEVLARAVQEA-----LVREKAALB--VRLQAVERDRODLAQLQ 760
    :      |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
783 QY HGAIAKNLPENENNTEGLN-VIN-----KVATFKPGVSGHYVELKDLSKDF 831
    :      |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
761 DB GLSSAKELLESSLFEAQQN$VIDEPQGLEVOIQTVTQAKEVIQGE-VRLCKLE----814
    :      |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
832 QY NMVFYHYSQTQS$AEHMKGRKOENKDEALPPPPPFCFPAFSKVINLLNCIDIMMYIL 891
    :      |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
815 DB LTERSQAQERDAARAQLAQAESQGTALBQQAAHEKVNQ-----L 858
    :      |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
892 QY RTVFERRAIDTSNLMT$EGMLQAFHIALGLLEBKQOLQAP$EETFDFTYHKASRLGSS 951
    :      |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
859 DB REKWK-----ERSWHQQLAKALESLEREKMELEMRLEKEQOTEMEAIQAOREERTQAE 913
    :      |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
952 QY ANNIQMILLEKLGIPOLEGQOMITWILQMFDTVKLR$KSLCI--VATTSSGESIKNDE 1009
    :      |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
914 DB SALCOM-----QULETEKERVSLLETILTQKELDASQOLERLRQDMKYQKLKEQB 964
    :      |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
1010 QY IT-----HDKEBAERKKAARLHRQKINA---QMSALOKNFLETHKLAYDNLTSEMPGKE 1062
    :      |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
965 DB TTGILTOLOEAQRELK-EARQRHDDLAALQEBSSLLQDKMQLQVEDUKSULVAQD 1023
    :      |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
1063 QY DS--IME-----BESTPAVSYSRIALGPGRGPSVTEKEVLTICLBQEBQVKIENNAWLVS 1117
    :      |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
1024 DB D$ORLVEQVBQEKLR$ETSYNRIQEKELR-----EKASLTLSLMEKEQRLLVLEQAD--S 1076
    :      |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
1118 QY ACVKORKSTALTOH-----RGKPIELSGE-----ALDPLFMDPDLAYTTGSCGHV 1162
    :      |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
1077 DB IRQOELSALURQMWQEAQGQSKELSAQWELLROEVKEKEADFLAQEAQLLEEL---EASHI 1133
    :      |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
1163 QY ----MHAVCWQKYFAVOL$SQORIHVLDLFDLESGYLCPCLKSLCNTVIPTIIPQPKI 1218
    :      |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
1134 DB TBQOLRASLWAQEAQAOL-----HLRLRSTES-----QLEALAAEQO 1171
    :      |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
1219 QY NGENADALAQLTLARIWIQTVLARI 1243
    :      |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
1172 DB PGNQAOQAQALASLYSALQOALGSV 1196
    :      |   |   |   |   |   |   |   |   |   |   |   |   |   |   |

```

Search completed: July 9, 2005, 15:32:46
Job time : 64 secs

THIS PAGE BLANK (USPTO)